

DIPLOMATURA D'ESTADÍSTICA

Títol: **Progressió del nombre d'arítmies**

Autor: **Raquel López Blázquez
Fabiana Montes de Oca Galizzi**

Director: **Erik Cobo Valeri**

Data: **Juny 2004**

UNIVERSITAT POLITÈCNICA DE CATALUNYA
Biblioteca



1400708258



Facultat de Matemàtiques
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA

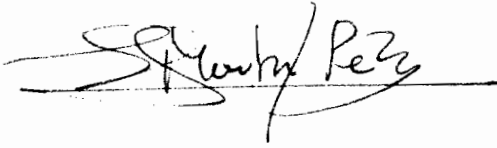
DADES DEL PROJECTE:

Nom de l'estudiant: Raquel López B. Laíquez
Feliciana Montes de Coca Grellizzi
DNI: 46748747M
X3654383J
Títol del Projecte: Progressió del nombre d'arbitres
Director del Projecte: ERK Cobo Velez
Tutor del Projecte:


QUALIFICACIÓ

Matricula d'honor (10)

MEMBRES DEL TRIBUNAL (nom i signatura)

President: Marta Pérez 

Vocal: F. Javier Heredia 

Secretari: ERK Cobo 

Data: 30 de juny de 2004

AGRAÏMENTS

A Joaquina Belchi, Xavier Navarro i Aurelio Quesada per la confiança que han dipositat en nosaltres, per haver-nos facilitat les dades i per estar disposats a ajudar-nos a entendre aspectes mèdics que fins ara desconexíem totalment.

A Javier Heredia per orientar-nos en els temes d'Investigació Operativa, a la Marta Pérez per assessorar-nos amb la sobredispersió i als altres profes que han col·laborat ens temes puntuals. A tots ells, gràcies pel bon tracte personal que ens han donat.

A Erik Cobo per haver-nos seguit tant de prop i per la paciència que ha tingut amb nosaltres.

A la Cris, molt especialment, per estar amb nosaltres des de primera hora del matí fins a la última de la nit, per orientar-nos en tots els aspectes, tant acadèmics com personals, per fer-nos riure, per calmar-nos ens moments d'estrès, pels cafès compartits i, en definitiva, per tot.

A la Raquel Sabarich per compartir amb nosaltres unes nits tant llargues i tantes confidències.

A la Helena Pérez i a Lula per ajudar-nos a solucionar alguns problemets d'última hora i per acompanyar-nos en moments de desconexió.

A Olga i Mireia per estar amb nosaltres des del principi i per compartir els caps de setmana a la facultat.

Al Tete Ramon per aguantar els nostres moments de davallada i alegrar-nos els dies.

Al grup de gent que ens ha donat tants moments macos i, esperem, que repetibles: Anna Brichs, Elena, Jordi Jimmy Jimeno, Èrica, Francesc, Miquel, Josep, Leila, Marc i Irene.

I a tots aquells que ens han acompanyat d'una manera o altre al llarg del nostre projecte: Francesc, Guillem, Esther, David, Misha, Oriol i Ruth.

Fabiana: Personalmente quiero agradecer a mi familia por haberme apoyado siempre y por haberme alentado en momentos de bajón. A Mario por su paciencia, sus consejos y por haberme dedicado tantas horas de largas conversaciones.

A mi “segunda familia” por haberme acogido tantas noches en su casa como una hija más. Y muy especialmente a mi “compi” Raquel por haberme hecho reír tanto a lo largo de este tiempo, por haber compartido horas de agobio, tristezas y alegrías, por haber sido tantas veces mi confidente y mi gran apoyo.

De verdad, MUCHAS GRACIAS A TODOS!!!

Raquel: L'agraïment amb més afecte cap a la meva família, ja que sense tots ells no hauria tirat endavant. També a tots aquells que han estat al meu costat i m'han donat ànims, molt especialment a la Patri i al David.

Per últim, donar-li les gràcies a la Fabi, no només per ser la meva companya de classe, projecte i pràctiques, sinó que també per haver-me demostrat que és una gran amiga amb la que sempre hi podré comptar. Gràcies per suportar-me, per compartir tant bons moments i pel *mate*!!!

RESUM

Introducció. Una arítmia es dona quan es produeix qualsevol irregularitat en el ritme natural de cor. A partir de les dades que tenim, on queda constància dels diferents episodis que han patit 59 pacients que tenen implantat un desfibril·lador automàtic (DAI), estudiarem la progressió del nombre d'arítmies.

Mètodes. Hem estudiat el temps entre arítmies consecutives, prèvia transformació logarítmica, interpretant els gràfics de Bland-Altman i realitzant els test de la t-Student per a dades aparellades. Tanmateix, hem construït un model de cadenes de Markov per estudiar l'evolució en quant al nombre d' arítmies i estimar les probabilitats de transició a 1, 12 i 36 mesos. Hem construït els intervals de confiança d'aquestes estimacions amb el mètode de remostreig del Jackknife.

Resultats. El procés d'aparició d'arítmies és més ràpid després del primer episodi, ja que el segon triga en aparèixer, amb una confiança del 95%, entre un 34'37% i un 57'87% menys que el primer, i que un cop s'ha patit el segon, aquest procés s'estabilitza. Amb el model de Markov, hem trobat que aquells pacients que tenen un major nombre d'arítmies en un mes determinat, tenen una major probabilitat de tenir-ne més el mes següent. Per exemple, si un pacient no ha tingut cap arítmia en un mes, la probabilitat de patir una o més d'una el mes següent és del 21'40% (11'50% ; 31'30%), si s'ha tingut una és del 40% (13'50% ; 66'50%) i si s'han patit dues és del 67% (0% ; 100%).

Discussió. Hem de tenir present que els nostres resultats han estat obtinguts a partir d'un nombre molt reduït de dades, i que, per tant, els estimadors calculats tenen intervals de confiança molt amples.

CAPÍTOL 1: Introducció

1.1. Motivació del projecte	1
1.2. Idea original	1
1.3. Importància del tema	2
1.4. Objectius	3

CAPÍTOL 2: Material i mètodes

2.1. La base de dades	4
2.2. Les arítmies	5
2.2.1. Categories d'arítmies	6
2.2.2. Com es tracta l'arítmia?	9
2.3. El desfibril·lador automàtic implantable (DAI)	10
2.3.1. Record històric	10
2.3.2. Aspectes tècnics	12
2.3.3. Teràpies programades	13
2.4. Variables	14
2.5. Plantejament de l'anàlisi estadístic	15
2.5.1. Cadenes de Markov	15
2.5.2. Metodologia per a la construcció de la cadena de Markov	20
2.5.3. Mètode de Jackknife	26
2.5.4. Intervalls de confiança amb la Distribució Binomial	28

CAPÍTOL 3: Resultats

3.1. Nombre d'episodis per pacient	30
3.2. Evolució del temps entre arítmies	34
3.2.1. Temps entre arítmies	35
3.2.2. Estudi de l'additivitat	
del temps entre arítmies	37
3.2.3. Estudi de l'additivitat en l'escala logarítmica	
del temps entre arítmies	40
3.3. Modelització del nombre d'arítmies	45
3.3.1. Matriu P de probabilitats de transició	45
3.3.2. Diagrama de la cadena de Markov	47
3.3.3. Estimacions per intervals de	
les probabilitats de transició	48
3.3.4. Estimacions de les probabilitats	
amb la Binomial	50
3.3.5. Probabilitats en m passos $P^{(m)}$	51
3.3.6. Taula resum	54

CAPÍTOL 4: Conclusions i discussions

4.1. Conclusions de l'estudi	55
4.1.1. Interpretació de les probabilitat de tenir arítmies	55
4.1.2. Evolució del temps entre episodis	56
4.2. Discussions	57
4.3. Estudis futurs	59

BIBLIOGRAFIA	63
---------------------------	-----------

ANNEXES

Annex 1: Article	65
Annex 2: Càlculs de probabilitats	71
Annex 3: Macro Minitab Jackknife	74
Annex 4: Estudi complementari de	
<i>l'evolució del temps entre arítmies</i>	75
Annex 5: Programa SAS Simulacont	83
Annex 6: Programa SAS Simuladisc	83
Annex 7: Col·laboració en	
<i>la tesi doctoral de Joaquina Belchi</i>	84

CAPÍTOL 1

INTRODUCCIÓ

1.1. MOTIVACIÓ DE PROJECTE

El nostre interès s'ha centrat en mostrar el potencial d'utilitzar simultàniament eines procedents de la Investigació Operativa i de l'Estadística. En aquest projecte fem ús de les cadenes de Markov i de l'estimació pel mètode de Jackknife per a resoldre un problema mèdic.

1.2. IDEA ORIGINAL

La idea de construir el model de cadenes de Markov per a representar les aritmies i de fer les estimacions per interval mitjançant un mètode de remostreig, ha sorgit de l'article, inclòs a l'Annex 1: *An application of Markov models in estimating transition probabilities for postmenopausal women with osteoporosis*. En ell es crea un model de Markov per a estudiar la progressió de les fractures vertebrals en dones amb osteoporosis menopausica. Tot i que la idea està extreta d'aquest estudi, hem hagut de fer una adaptació per tal de què el mètode sigui adequat per a les nostres dades.

1.3. IMPORTÀNCIA DEL TEMA

En els problemes de presa de decisions ens trobem sovint problemes basats en fenòmens que tenen algun tipus d'incertesa associada. Aquesta incertesa o aleatorietat pot estar causada per fonts de variabilitat que s'escapen del control o per la complexitat dels fenòmens naturals. En comptes de tractar la variabilitat de manera qualitativa, pot incorporar-se al model matemàtic i rebre un tractament quantitatiu, sempre que el fenomen presenti algun grau de regularitat, de manera que la seva variabilitat pugui ser descrita per un model probabilista. Les cadenes de Markov han estat utilitzades en nombroses recerques ja que són una eina molt útil per a modelar aspectes de la vida real.

En el nostre projecte, fem servir aquests models per tal d'observar si la probabilitat de tenir una arítmia depèn del nombre d'arítmies prèvies.

En aquest estudi, mostrarem com modelar la progressió del nombre d'arítmies que pateix un pacient usant les cadenes de Markov. Així doncs, estimarem les probabilitats de transició a un mes i d'aquestes en derivaran les matrius de probabilitats a m mesos. Per tal de construir els intervals de confiança de totes aquestes probabilitats estimades, s'ha proposat utilitzar el mètode de Jackknife ja que preserva l'estructura de correlacions entre les unitats de temps utilitzades per a estimar les probabilitats d'una cadena de Markov. Per una altra banda, calcularem les estimacions dels intervals de confiança de les probabilitats de transició amb la distribució Binomial per comparar l'amplada dels intervals.

Un altre punt important del projecte serà l'estudi del temps entre arítmies. Aquí veurem com progressa el temps de desenvolupament atípic del ritme cardíac dels pacients. Així podrem observar si existeix una acceleració, una desacceleració o una permanència constant dels episodis en el temps.

1.4. OBJECTIUS

Com veurem més endavant, aquest projecte tindrà tres anàlisis importants, pels quals ens hem plantejat els següents propòsits: estudiar la variable que compta el nombre d'episodis de cada pacient; veurem si aquells pacients que tenen un major nombre d'arítmies en un mes determinat, tenen una major probabilitat de tenir-ne més el mes següent, mitjançant un model de cadenes de Markov; estudiar l'evolució del temps entre arítmies i veure si el procés es va fent més ràpid a mesura que apareixen nous episodis. Així doncs, el nostre objectiu serà obtenir informació que ens pugui ajudar a entendre la progressió de les arítmies en pacients amb Desfibril·lador Automàtic Implantable (DAI) per tal poder predir el seu desenvolupament al llarg del temps.

CAPÍTOL 2 MATERIAL I MÈTODES

2.1. LA BASE DE DADES

La tesis doctoral de la Dra. Joaquina Belchi consisteix en implantar un dispositiu (DAI) en dues poblacions diferents de malalts de cor (veure *Annex 7* on expliquem la nostra col·laboració en la seva tesi). Per una banda, pacients que pateixen arítmies malignes que posen en perill la vida, les *ventriculars*, però que a la vegada per característiques del cor, poden desenvolupar també arítmies benignes, les *auriculars*. Per una altra banda, pacients que pateixen només arítmies benignes, però que resisteixen als fàrmacs habituals i, encara que és poc freqüent, poden desenvolupar arítmies malignes.

El dispositiu implantat en ambdós casos és el mateix, amb idèntica capacitat per diagnosticar i tractar les arítmies, siguin malignes o benignes. La forma en què es programa el dispositiu es deu al criteri del metge.

Al llarg d'un seguiment mig de 56 ($\pm 16'41$) mesos, un total de 4766 episodis d'arítmies espontànies de 59 pacients han estat documentades pel dispositiu DAI.

Per portar a terme el seu estudi, ha disposat de dues bases de dades en SPSS: la primera, denominada *pacients*, formada per 59 observacions i 37 variables, que fa referència a les característiques demogràfiques (edat, sexe, ...) i altres variables clíniques, com per exemple, cardiopatia base on es classifiquen els malalts en funció de la indicació. La segona base de dades, denominada *episodis*, on es té constància de les 64 variables i dels 4766 episodis que han tingut els pacients en global.

2.2. LES ARÍTMIES

El cor batega gairebé 5 litres de sang per l'organisme cada 60 segons. Fins i tot en repòs, el cor es dilata i es contrau entre 60 i 100 vegades per minut. Aquests batecs estan provocats per impulsos elèctrics que s'originen en el marcapassos natural del cor, el nòdul sinodal o sinoauricular (nòdul SA). El nòdul SA és un grup de cèl·lules situades a la part superior de la cavitat superior dreta del cor (la aurícula dreta).

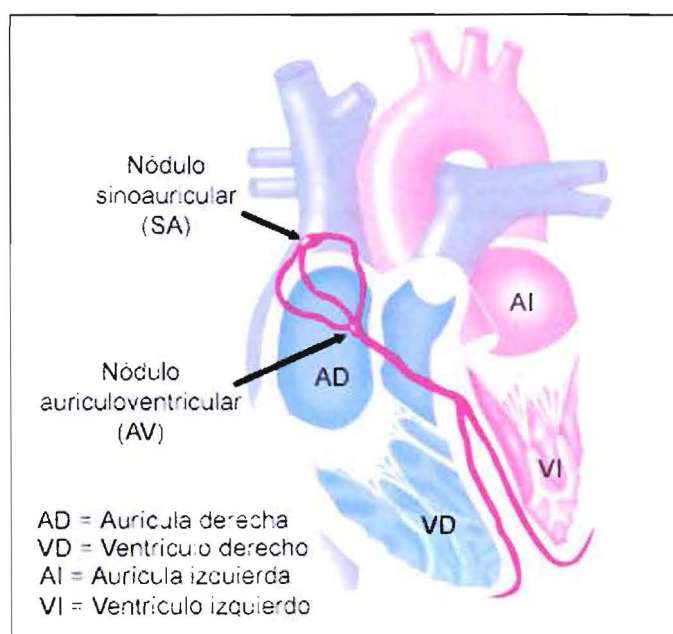


Fig.1. Parts del cor

Tota irregularitat en el ritme natural del cor s'anomena *arítmia*. Qualsevol pot sentir batecs irregulars en algun moment de la seva vida; palpitations lleus i poc freqüents, que són inofensives.

2.2.1. Categories d'arítmies

Les arítmies poden dividir-se en dues categories: ventriculars i supraventriculars. Les arítmies ventriculars es produeixen en les cavitats inferiors del cor anomenades ventricles. Les arítmies supraventriculars es produeixen a la zona que es troba per sobre dels ventricles, generalment a les aurícules, que són les cavitats superiors del cor.

Les arítmies també es defineixen segons la velocitat dels batecs. La bradicàrdia és un pols molt lent mentre que la taquicàrdia és un pols molt ràpid. El tipus més greu d'arítmia és la fibril·lació, que consisteix en contraccions irregulars de fibres musculars cardíques individuals de manera no coordinada.

Bradicàrdia

La bradicàrdia és una freqüència cardíaca molt baixa de menys de 60 batecs per minut. Es produeix quan l'impuls elèctric que estimula la contracció del cor no es genera al marcapassos natural del cor, el nòdul sinodal o sinoauricular (nòdul SA), o no és enviat a les cavitats inferiors del cor (els ventricles) per les vies correctes. La bradicàrdia afecta principalment a les persones majors, però pot afectar a persones de qualsevol edat, fins i tot a nens molt petits. Pot tenir una de dues causes: el sistema nerviós central no comunica al cor que deu bombar més o el nòdul SA podria estar danyat. Aquest mal pot deure's a una malaltia cardiovascular, al procés d'envelliment o a defectes heretats o congènits. També podria ser causat per certs medicaments, fins i tot aquells que s'administren per a controlar les arítmies i la pressió arterial alta.

Taquicàrdia

La taquicàrdia és una freqüència cardíaca molt elevada de més de 100 batecs per minut. Hi ha molts tipus diferents de taquicàrdia, segons on s'origini el ritme accelerat. Si s'origina en els ventricles, es denomina *taquicàrdia ventricular*. Si s'origina per sobre dels ventricles, es denomina *taquicàrdia supraventricular*.

Arítmies ventriculars

Taquicàrdia ventricular

La taquicàrdia ventricular és quan el nòdul SA ja no controla el batec dels ventricles, sinó que altres zones al llarg de la via de conducció elèctrica inferior assumeixen la funció de marcapassos. Com el nou senyal no es desplaça pel múscul cardíac per la via normal, el múscul cardíac no batega en forma habitual. S'acceleren els batecs del cor i el pacient sent palpitations. Aquest ritme irregular pot produir una extrema falta d'alè, mareig o desmai.

Fibril·lació ventricular

L'arítmia més greu és la fibril·lació ventricular, que són batecs irregulars no controlats. En lloc de tenir una sola pulsació fora de temps als ventricles, és possible que diversos impulsos s'originin al mateix temps en diferents llocs, tots ells estimulant al cor a bategar. Per consegüent, es produeixen batecs molt més ràpids i desordenats que poden arribar als 300 batecs per minut. A causa d'aquests batecs caòtics, el cor bomba molt poca sang al cervell i a la resta de l'organisme, i és possible que la persona es desmai. És necessari obtenir assistència mèdica immediatament. Si és possible iniciar mesures de reanimació cardiopulmonar (RCP) o administrar descàrregues elèctriques per a restablir el ritme normal del cor, és possible que el cor no sofreixi un mal molt greu. Les persones que sofreixen d'alguna malaltia cardiovascular o que tenen antecedents d'atacs cardíacs tenen el major risc de patir fibril·lació ventricular.

Arítmies supraventriculars

Les arítmies supraventriculars es produeixen en les cavitats superiors del cor. En general, les arítmies supraventriculars, també denominades *arítmies auriculars*, no són tan greus com les ventriculars. En alguns casos, ni tan sols necessiten tractament. Les arítmies auriculars poden ser causades per diversos factors, com per exemple, el tabac, l'alcohol, la cafeïna i els medicaments per a la tos i els refredats. El trastorn també pot deure's a la cardiopatia reumàtica o a una tiroïdes hiperactiva (hipertiroïdisme).

Taquicàrdia supraventricular

La taquicàrdia supraventricular és una freqüència cardíaca regular però elevada, superior als 150 batecs per minut, que s'origina a les aurícules. A diferència d'altres tipus d'arítmia, la taquicàrdia supraventricular no s'origina al nòdul SA.

Fibril·lació auricular

La fibril·lació auricular és un ritme ràpid i irregular a causa de moviments convulsius de fibres musculars individuals del cor. És una de les principals causes d'accidents cerebrovasculars, especialment en les persones majors. La fibril·lació auricular pot fer que s'acumuli sang en les cavitats superiors del cor. La sang acumulada pot formar grumolls denominats *coàguls*. Si un coàgul de sang es desplaça del cor al cervell i obstrueix una de les artèries cerebrals que són més petites, pot produir-se un accident cerebrovascular. Aproximadament el 15% dels accidents cerebrovasculars es produeixen en persones amb fibril·lació auricular.

Bloqueig cardíac

Es produeix un bloqueig cardíac quan el nòdul SA envia correctament el senyal elèctric però aquest no passa pel nòdul auriculoventricular (AV) o les vies de conducció elèctrica inferiors amb la rapidesa deguda. Els bloqueigs es deuen principalment al procés d'envelliment o a una inflamació o cicatrització del cor que de vegades és ocasionada per la malaltia arterial coronària. Existeixen diverses classes diferents de bloqueig cardíac i es classifiquen segons la seva gravetat.

El bloqueig cardíac de primer grau és quan els impulsos que passen pel nòdul AV són massa lents. El bloqueig cardíac de segon grau és quan els impulsos passen per les aurícules del cor però es retarden al nòdul AV. A causa d'aquest retard, els ventricles no bateguen en el moment degut. El bloqueig cardíac de tercer grau és quan no arriben impulsos als ventricles. Per a compensar, els ventricles utilitzen el seu propi marcapassos *auxiliar* amb la seva freqüència més baixa. Com és probable que transcorri un espai de temps entre l'impuls de les aurícules i el de la unitat auxiliar, la persona pot desmaiar-se. El bloqueig cardíac de tercer grau és molt greu i pot donar lloc a insuficiència cardíaca i fins i tot la mort.

2.2.2. Com es tracta l'arítmia?

Sovint la primera mesura que es pren per a tractar l'arítmia és l'administració de medicaments antiarítmics, tal com la digital, els betabloquejants i els bloquejants càlcics. Altres tractaments inclouen les intervencions transcatèter, els dispositius implantables i la cirurgia (en casos extrems).

La taquicàrdia ventricular i la fibril·lació ventricular poden tractar-se mitjançant la implantació d'un desfibril·lador cardioversor, un dispositiu que administra impulsos elèctrics o, de ser necessari, una descàrrega, per a restablir el ritme normal del cor. El generador de l'aparell s'implanta en una butxaca sota la pell del pit o de l'abdomen i es connecta a pegats que es col·loquen sobre el cor. Els dispositius implantables més moderns s'introdueixen pels vasos sanguinis, sense necessitat de realitzar una operació de tòrax obert.

En alguns casos de freqüència cardíaca baixa, s'utilitza un marcapassos electrònic. Més petit que una caixa de fòsfors, el marcapassos electrònic s'implanta quirúrgicament prop de la clavícula, l'os que es troba sota el coll.

L'ablació per radiofreqüència és un procediment en el qual s'utilitza un catèter i un dispositiu que permet obtenir un mapa de les vies de conducció elèctrica del cor. Després d'administrar un relaxant al pacient, s'introdueix un catèter per una vena fins arribar al cor. Utilitzant ones electromagnètiques d'alta freqüència, els metges poden destruir les vies de conducció responsables de l'arítmia.

En alguns casos, aquests tractaments poden no ser eficaços o apropiats i podria ser necessari realitzar una intervenció quirúrgica per a destruir la zona on s'originen els batecs irregulars.

L'ablació quirúrgica és com l'ablació per radiofreqüència. Utilitzant tècniques de cartografia per computadora, els cirurgians poden descobrir les cèl·lules on s'originen els batecs irregulars. A continuació, amb una tècnica denominada crioablació, poden eliminar el teixit amb un escandall fred i destruir les cèl·lules defectuoses.

La cirurgia de Coix (tècnica de laberint) podria ser indicada en casos de fibril·lació auricular que no responen a medicaments o descàrregues elèctriques (tractament amb cardioversió). Els cirurgians creen un "laberint" de noves vies de

conducció elèctrica perquè els impulsos elèctrics puguin travessar el cor sense ser bloquejats.

La resecció ventricular permet que el cirurgià extirpi la zona del múscul cardíac on s'origina l'arítmia.

En alguns casos, no es necessita cap tractament i la majoria de les persones que sofreixen d'arítmia duen una vida normal i activa. Sovint, certs canvis en l'estil de vida, com per exemple evitar la cafeïna (en coses com el cafè, el te, les begudes gasoses, la xocolata i alguns analgèsics) o l'alcohol, són suficients per a eliminar l'arítmia.

2.3. EL DESFRIBIL·LADOR AUTOMÀTIC IMPLANTABLE (DAI)

2.3.1. Record històric

Com hem vist, una manera de tractar una arítmia és amb la implantació d'un desfibril·lador cardioversor (ICD).

La història del desfibril·lador implantable abasta poc més d'un quart de segle. El pioner en el desenvolupament inicial va ser el Dr. Michel Mirowski, que en 1966 i afectat per la mort del seu professor i amic el Dr. Harry Heller per episodis recurrents de taquicàrdia ventricular, va concebre la idea d'un sistema capaç de ser transportat i amb eficàcia en el sensat, cardioversió i desfibril·lació:

“Com haguéssim pogut prevenir en aquest moment la mort del Dr. Heller; mantenir-lo eternament ingressat en la Unitat Coronària, o seguir-lo a totes parts amb un desfibril·lador extern? Ambdues solucions eren òbviament impossibles. Així doncs, reflexionava, necessitàvem crear un aparell implantable com un marcapassos que vigilés la fibril·lació ventricular i desfibril·lés automàticament al pacient fins arribar al ritme sinodal”.

Durant aquest període el Dr. Mirowski vivia a Israel i, davant la falta de mitjans humans i financers per a poder desenvolupar el seu projecte, va decidir acceptar al 1969 el lloc de cap de la Unitat Coronària a l'Hospital Sinaí de Baltimore. A la fi de la dècada dels 60 i gràcies al suport del Dr. Mendeloff, per aquella època director de

l'Hospital, el Dr. Mirowski al costat del Dr. Morton Mower, van construir el primer prototip de desfibril·lador pel seu ús en animals. Aquest dispositiu disposava de les funcions bàsiques: detecció apropiada de fibril·lació ventricular i desfibril·lació interna amb èxit, utilitzant descàrregues de més de 20 Joules.

En 1972, i gràcies a la col·laboració financera de Medrac Inc, una empresa fabricadora de bombes d'injecció (Intec Systems) de Pittsburgh, Pennsylvania, va desenvolupar el primer sistema ICD clínic.

En 1976 ja es tenia el prototip d'implant crònic en animals i es va procedir a provar el sistema. Per a això, es va implantar el dispositiu en un gos que, a més, duia un fibril·lador implantable.

La primera experiència en humans es va portar a terme el 4 de Febrer de 1980 en l'Hospital Johns Hopkins, a una dona californiana amb múltiples episodis de mort sobtada. L'èxit d'aquesta primera experiència en humans va fer que el model original s'implantés en 37 pacients. Un segon prototip va incorporar la possibilitat de cardioversió.

El primer dispositiu que es va implantar en humans era un sistema gran, pesat, que requeria implant abdominal i toracotomia oberta per a implantar els pegats de desfibril·lació i sensat. Des de llavors, tant els investigadors com companyies comercials s'han esforçat a resoldre les limitacions inicials, fins arribar als dispositius actuals.

Les tècniques quirúrgiques han evolucionat de forma paral·lela al desenvolupament tecnològic del dispositiu. Així, la disminució de la grandària (de 168cc a 30cc) i del pes (de 290gr a 70gr) del generador han permès poder implantar-lo a nivell pectoral.

El primer implant d'un sistema completament endocàrdic es va realitzar en 1986 i el seu principal avantatge va ser l'eliminació de la toracotomia. L'inconvenient va ser que en un 20-40% dels pacients s'obtenien llindars de desfibril·lació massa alts. Amb la introducció de l'ona bifàsica en 1992, es va fer possible la implantació endovenosa en el 100% dels pacients, gràcies a una disminució significativa del llindar de desfibril·lació.

2.3.2. Aspectes tècnics

El DAI està constituït bàsicament per un generador implantable i un o més elèctrodes (Fig. 2). El generador conté les bateries, el generador d'impulsos, circuits, memòries i programes lògics necessaris pel correcte funcionament de les capacitats diagnòstiques i terapèutiques del sistema. Els elèctrodes permeten la detecció del senyal cardíac, l'estimulació elèctrica i la transmissió d'energia de desfibril·lació. A través del senyal recollit dels elèctrodes, s'identifica el ritme cardíac i si en algun moment es compleixen els requisits preprogramats de detecció d'una arítmia, el generador envia el tractament previst, de forma esglaonada.

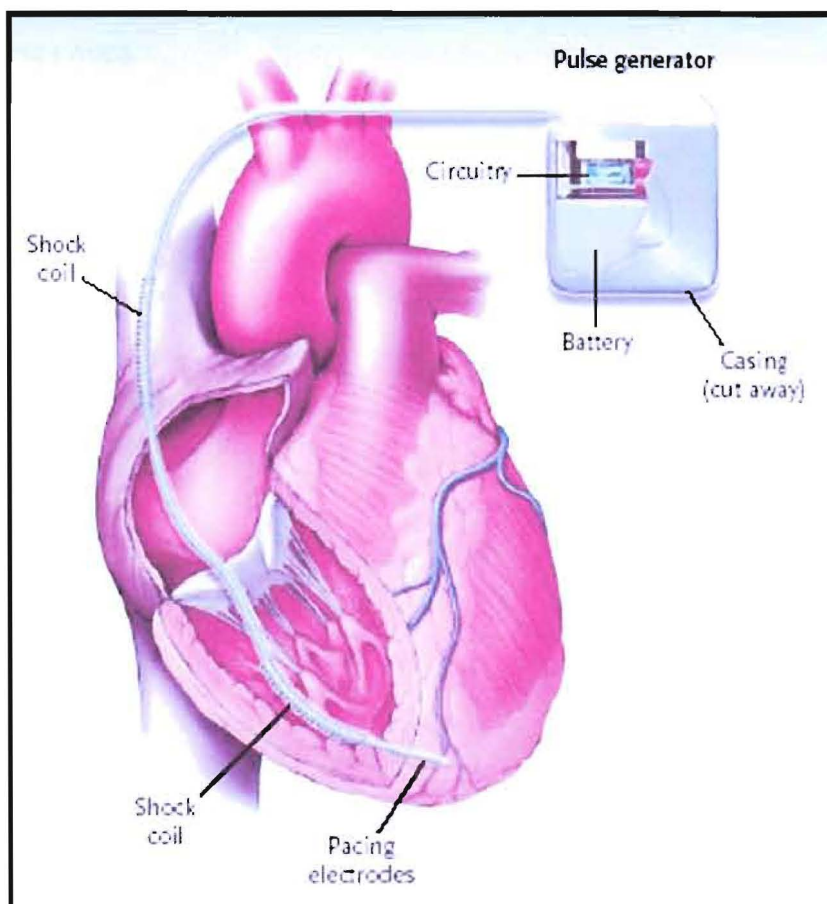


Fig.2. Constituents del DAI

2.3.3. Teràpies programades

Una vegada implantat el dispositiu, aquest comença a rebre informació del cor, detectant si el senyal cardíac és correcte. En cas de què el dispositiu identifiqui que el comportament del ritme cardíac és atípic, el programa començarà a treballar identificant el tipus d'arítmia i tot seguit aplicarà la teràpia corresponent: Rampa, Burst, 50 Hz o Xocs. Una vegada finalitzada la teràpia també es tenen en compte el nombre de seqüències de la teràpia, el nombre de teràpies que s'ha subministrat al malalt, i el resultat d'aquestes. Si la teràpia no és eficaç es pot subministrar novament la mateixa o canviar-la per una altra, ja que els tipus d'arítmies poden anar canviant en un mateix episodi per un mateix pacient. Pel que fa als episodis TA, la seqüència de teràpies serà: Burst, Rampa, 50 Hz i per últim Xocs, en canvi per als episodis FA la seqüència serà: 50 Hz i Xocs.

2.4. VARIABLES

Com podrem veure, en el nostre projecte hem realitzat diversos enfocaments.

Per portar a terme l'estudi del nombre d'episodis per pacient, hem hagut d'utilitzar la variable *pacients* de la base de dades *episodis* per calcular les freqüències d'arítmies per a cada malalt, creant així la variable *numepi*, que compta els episodis que ha tingut cada pacient durant el temps que ha estat en seguiment. Aquesta variable l'hem afegit a la base de dades de *pacients*. Tot seguit, hem dividit *numepi* pel nombre de mesos de seguiment, variable *t_seg_m* de la base de dades *pacients*, de cada malalt i hem creat la variable *taxaepi*. Aquesta última l'hem afegit a la base de dades de *pacients*.

Per fer l'estudi de l'evolució dels temps entre arítmies, hem hagut de crear les variables t_{01} , t_{12} i t_{23} , on t_{ij} és el temps que passa entre la arítmia *i*-èssima i la *j*-èssima, a partir de les variables temps fins la primera arítmia, temps fins la segona i temps fins la tercera des de l'implant del DAI. Així, fent la diferència obtenim el temps entre una arítmia i la següent. Cal destacar que en el cas de t_{01} tenim en compte el temps que ha passat des de què es fa l'implant fins a l'aparició de la primera arítmia. En aquest cas t_{01} està censurada per l'esquerra ja que no es té registrat el temps que ha passat des de l'últim episodi abans de l'implant.

Pel que fa a l'estudi amb Cadenes de Markov, hem comptabilitzat el nombre d'episodis que cada pacient ha patit en els tres primers mesos i s'han calculat els seus acumulats. Així s'obtenen les variables m_1 , m_2 i m_3 tals que m_i és el nombre d'arítmies acumulades fins al final del mes *i*.

2.5. PLANTEJAMENT DE L'ANÀLISI ESTADÍSTIC

El pla estadístic del nostre estudi engloba diferents metodologies. Per fer una estimació de les probabilitats de tenir arítmies en un mes, coneixent les que s'ha tingut en el mes anterior, es farà servir el model de cadenes de Markov, que ens proporciona una estructura matricial on queden especificats aquests aspectes. Per tal de fer un anàlisi de com es comportarà aquest procés un cop hagin passat un i tres anys, s'estudiaran les probabilitats a llarg termini.

D'altra banda, cal tenir present que aquests mètodes ens proporcionen les estimacions puntuals de les probabilitats i que, per tal de considerar les oscil·lacions aleatòries d'aquests estimadors, ens caldrà construir els intervals de confiança. Per fer-ho, es farà servir el mètode de Jackknife que preserva l'estructura de correlació entre les probabilitats de transició d'una cadena de Markov (les estimacions de les probabilitats no són independents com veurem a continuació).

2.5.1. Cadenes de Markov

La següent explicació teòrica sobre les cadenes de Markov ha estat extreta dels apunts del professor de l'assignatura d'Investigació Operativa Estocàstica, Esteve Codina.

Procés estocàstic

Un *procés estocàstic* és una col·lecció indexada de variables aleatòries $\{X_t \mid t \in T\}$, on l'índex t pertany a un conjunt T conegut.

Els valors de les variables aleatòries X_t s'anomenen estats i els seus valors possibles estan dins de l'anomenat espai d'estats, el qual es notará per I . Si l'espai d'estats d'un procés estocàstic és discret aleshores s'anomena cadena i en aquest cas $I = \{0, 1, 2, 3, \dots\}$.

Poden haver-hi molts processos estocàstics d'interès en un sistema de característiques no deterministes a estudiar, alguns més assequibles de tractar matemàticament que d'altres. En general, els processos estocàstics que s'analitzen responen a la següent estructura: el sistema es troba en exactament una de les possibles categories o estats que defineixen l'espai d'estats discret i finit $I = \{0,1,2,3,\dots\}$, al llarg d'instantants de temps t etiquetats com a $0,1,2,3,\dots$. Els instants poden espaiar-se equitativament en el temps (per exemple, cada dia, cada setmana, cada mes,...) o no, segons el sistema a estudiar (per exemple, en funció del comportament d'altres elements del sistema). Les etiquetes dels estats poden constituir una caracterització quantitativa o qualitativa del sistema modelat sense implicar cap pèrdua de la potència modelística.

El present projecte estudiarà un cas particular de les cadenes de paràmetre discret $\{X_t \mid t \in T\}$ on $T = \{0,1,2,3,\dots\}$ i espai d'estats $I = \{1,2,3,\dots,M\}$ (valors de les variables aleatòries X_t): les cadenes de Markov.

Cadenes de Markov

Les cadenes de Markov constitueixen la representació matemàtica d'un sistema físic a estudiar, on les variables aleatòries X_t que caracteritzen el sistema són observades en instantants de temps $t = 0,1,2,3,\dots$, i on cada variable aleatòria pot prendre un valor qualsevol entre els M valors dins de I que constitueixen una caracterització dels M estats del procés.

El tractament analític del processos estocàstics requereix de certes hipòtesis simplificadores pel què fa a la distribució conjunta de les variables aleatòries Y_0, Y_1, Y_2, \dots . Una de les hipòtesis simplificadores més utilitzada rep el nom de *propietat markoviana*.

Propietat Markoviana

Un procés estocàstic $\{X_t \mid t = 0, 1, 2, 3, \dots\}$ de paràmetre discret gaudeix de la *propietat markoviana* si:

$$P(\{X_{t+1} = j\} \mid \{X_t = i, X_{t-1} = x_{t-1}, X_{t-2} = x_{t-2}, \dots, X_1 = x_1, X_0 = x_0\}) = P(\{X_{t+1} = j\} \mid \{X_t = i\})$$

per tot valor $t = 0, 1, 2, 3, \dots$ i tota seqüència d'estats $i, j, x_{t-1}, x_{t-2}, \dots, x_1, x_0$.

La propietat markoviana pot veure's que és equivalent a afirmar que la probabilitat condicional de qualsevol esdeveniment en el futur, donat un esdeveniment qualsevol en el passat i l'estat actual $X_t = i$, és independent de la trajectòria en el passat i només depèn de l'estat actual del procés. Les probabilitats condicionals $P(\{X_{t+1} = j\} \mid \{X_t = i\})$ s'anomenen *probabilitats condicionals de transició* de l'estat i a l'estat j en un interval de temps o pas.

Concepte d'estacionarietat

Un altre hipòtesi simplificadora rau en l'*estacionarietat de les probabilitats de transició* i pot descriure's de la següent manera:

$$P(\{X_{t+1} = j\} \mid \{X_t = i\}) = P(\{X_t = j\} \mid \{X_0 = i\}) \quad \text{per a tot } t = 0, 1, 2, 3, \dots$$

L'estacionarietat implica que les probabilitats condicionals de transició entre estats no varien al llarg del temps.

Considerant les dues hipòtesis conjuntament es poden definir en el sistema probabilitats de transició estacionàries en un pas i construir una matriu P quadrada amb files i columnes pels M estats de la cadena on a cada cel·la (i, j) conté la probabilitat de transició en un pas de l'estat i a l'estat j , p_{ij} . És a dir,

$$P = \begin{bmatrix} p_{11} & \dots & p_{1M} \\ \dots & \dots & \dots \\ p_{M1} & \dots & p_{MM} \end{bmatrix} \quad \text{on } P(\{X_{t+1} = j\} \mid \{X_t = i\}) = p_{ij} \text{ per tot } t = 0, 1, 2, 3, \dots$$

La matriu P d'una cadena que gaudeix de les dues hipòtesis simplificadores és una matriu anomenada estocàstica ja que verifica:

1. $p_{ij} \geq 0$ per a tot i, j , les components són no negatives
2. $\sum_{j=1}^M p_{ij} = 1$ per qualsevol fila i , el seus elements sumen la unitat

Probabilitats de transició en m passos

Anant una mica més enllà, l'existència de probabilitats de transició estacionàries també implica que les probabilitats de transició en m passos també són estacionàries, és a dir:

$$P(\{X_{m+1} = j\} | \{X_t = i\}) = P(\{X_m = j\} | \{X_0 = i\}) \text{ per a tot } t = 0, 1, 2, 3, \dots$$

Aquestes probabilitats condicionals es noten,

$$p_{ij}^{(m)}, \quad \text{on } P(\{X_m = j\} | \{X_0 = i\}) = p_{ij}^{(m)}$$

i representa la probabilitat que el procés estigui a l'estat j després de m passos (unitats de temps) condicionat a haver començat a l'estat i . A la vegada defineixen una matriu quadrada que té com a dimensió el nombre d'estats de la cadena:

$$P^{(m)} = \begin{bmatrix} p_{11}^{(m)} & \dots & p_{1M}^{(m)} \\ \dots & \dots & \dots \\ p_{M1}^{(m)} & \dots & p_{MM}^{(m)} \end{bmatrix} \text{ per a tot } m = 0, 1, 2, 3, \dots$$

Per $m=1$ correspon a la matriu de transicions en un pas: $P^{(1)} = P$.

Per $m=0$ correspon $P^{(0)} = I_M$, que és la matriu identitat de dimensió M , doncs

$$P^{(0)} = P(\{X_0 = j\} | \{X_0 = i\}) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{altrament} \end{cases}$$

Per qualsevol nombre de passos m , $P^{(m)}$ és una matriu estocàstica, ja que:

1. $p_{ij}^{(m)} \geq 0$ per $m = 0, 1, 2, 3, \dots$ i per a tot i, j , les components són no negatives
2. $\sum_{j=1}^M p_{ij}^{(m)} = 1$ per $m = 0, 1, 2, 3, \dots$ i per qualsevol fila i , el seus elements sumen 1.

De manera més general es pot veure que $P^{(m)}$ es pot obtenir de multiplicar matricialment la matriu de probabilitats de transició P per la matriu de probabilitats condicionals de transició en $m-1$ passos, $P^{(m-1)}$:

$$P^{(m)} = P \cdot P \dots P = P^m = P^{(m-1)} \cdot P$$

O bé, considerant $l=m-1$:

$$P^{(m)} = P \cdot P \dots P = P^m = P \cdot P^{(m-1)}$$

És a dir, la matriu $P^{(m)}$ de probabilitats condicionals de transició en m passos és la potència m -èsima de la matriu de transició a un pas P i pot calcular-se iterativament. S'aconsella tenir cura dels errors numèrics que poden sorgir en els càlculs si m és força gran.

Cadena de Markov finita

Un procés estocàstic $\{X_t \mid t = 0, 1, 2, 3, \dots\}$ és una *cadena de Markov finita* si:

1. Té un nombre finit d'estats, M .
2. Compleix la propietat markoviana.
3. Les probabilitats de transició són estacionàries.
4. Té un conjunt de probabilitats inicials $P(\{X_0 = i\})$ per a tot estat i .

2.5.2. Metodologia per a la construcció de la cadena de Markov

A continuació presentem el procediment seguit per tal de fer la construcció de la cadena de Markov que ens modela el nombre d'arítmies dels pacients.

Modificacions de les variables

Per portar a terme aquest estudi ens caldrà fer algunes modificacions de les bases de dades originals, per tal de tenir les dades de la manera necessària per a construir el model de cadenes de Markov.

En primer lloc, hem realitzat una selecció de les arítmies que han succeït en els tres primers mesos després de l'implant del DAI. D'aquesta manera ens quedem amb la informació dels 59 pacients que hi ha a l'estudi, ja que tots han estat seguits més de tres mesos. A continuació, s'han comptabilitzat el nombre d'arítmies que cada pacient ha patit en el mes 1, en el mes 2 i en el mes 3, i s'han calculats els seus acumulats. Així s'obtenen les variables m_1 , m_2 i m_3 tals que:

m_1 : nombre d'arítmies acumulades fins al final del mes 1

m_2 : nombre d'arítmies acumulades fins al final del mes 2

m_3 : nombre d'arítmies acumulades fins al final del mes 3

Tot seguit, es presenta l'anàlisi descriptiu d'aquestes variables amb els gràfics corresponents:

Variable	N	Mean	Median	TrMean	StDev	SE Mean
m1	59	5,17	0,00	1,64	17,73	2,31
m2	59	6,20	0,00	1,74	22,18	2,89
m3	59	5,31	0,00	3,30	12,42	1,62

Variable	Minimum	Maximum	Q1	Q3
m1	0,00	104,00	0,00	1,00
m2	0,00	128,00	0,00	1,00
m3	0,00	50,00	0,00	2,00

Fig. 3. Descriptiva del nombre d'arítmies acumulades

Pel que fa a les arítmies acumulades fins al mes 1 tenim:

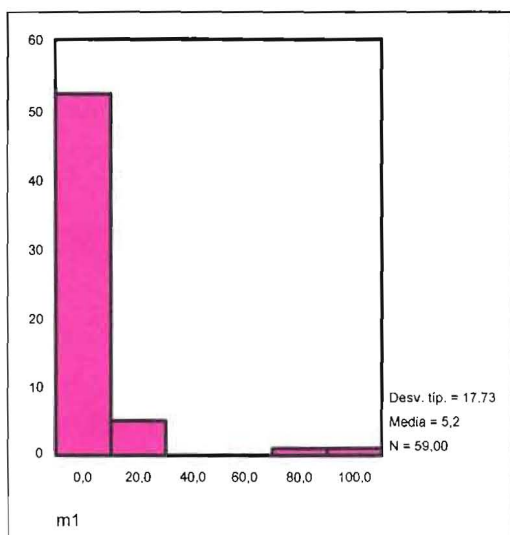


Fig. 4. Histograma de la variable m1

Com es pot observar a la Fig. 4., al final del mes 1, la major part dels pacients tenen entre 0 i 10 arítmies acumulades. El rang de valors de la variable va fins a 104 arítmies. Cal notar que els quartils 1 i 2 es troben a zero.

Pel que fa a les arítmies acumulades fins al mes 2 tenim:

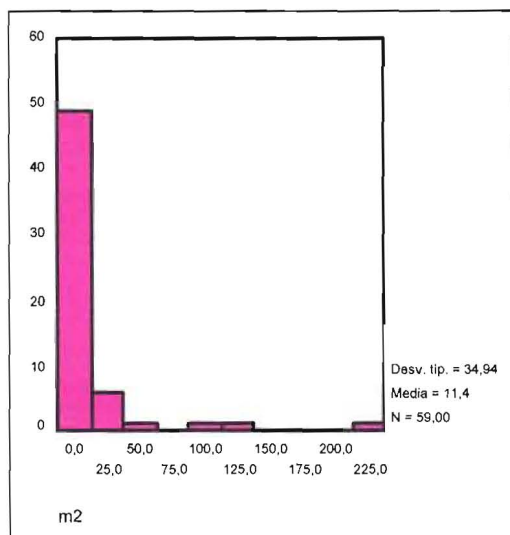


Fig. 5. Histograma de la variable m2

Aquest gràfic ens mostra que la majoria de pacients tenen menys de 25 arítmies acumulades al final del mes 2. El rang de valors de la variable ha augmentat fins ha 214 arítmies.

Pel que fa a les arítmies acumulades fins al mes 3 tenim:

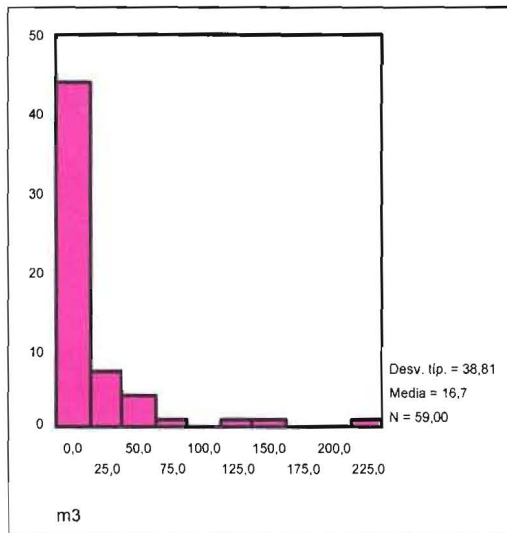


Fig. 6. Histograma de la variable m3

Aquesta variable és encara més dispersa però, tot i havent-hi una primera barra on s'acumulen la majoria de pacients, sembla que quedin més repartits per la resta de valors que en els dos gràfics anteriors (Fig. 4 i Fig. 5).

Ara bé, s'ha decidit discretitzar aquestes variables en sis categories que després seran els estats de la cadena de Markov: *estat 0*, correspon a tenir zero arítmies acumulades, *estat 1*, una arítmia acumulada, *estat 2*, dues arítmies acumulades, *estat 3*, de tres a sis arítmies acumulades, *estat 4*, de set a vint arítmies acumulades i *estat 5*, correspon a tenir més de vint arítmies. Les variables modificades amb la nova nomenclatura són n_1 , n_2 i n_3 , on n_i ens indica el nombre d'arítmies acumulades fins al mes i un cop discretitzada la m_i .

L'elecció de la nova codificació es basa, no només en la seva distribució empírica, sinó que també en els interessos propis de l'estudi, ja que és més important conèixer el comportament de les primeres arítmies que no pas el de les últimes (el pas de 0 a 1 arítmies és rellevant, però no ho és tant el de 100 a 101).

Configuracions

Amb aquestes tres variables hem construït la configuració de cada pacient (n_1, n_2, n_3) , on queda representada l'evolució del nombre d'arítmies de cadascun d'ells durant els tres primers mesos de l'estudi. Per exemple, el malalt que tingui al final del mes 1, zero arítmies, al final del mes 2, una i al final del mes 3, dues, tindrà la configuració $(0,1,2)$.

Com que estem comptabilitzant arítmies acumulades tenim la garantia de que $0 \leq n_1 \leq n_2 \leq n_3 \leq 5$, és a dir, no es podrà retrocedir d'estat. El rang queda definit entre 0 i 5 ja que, com hem comentat abans, l'espai d'estats de la cadena de Markov que construirem serà $I = \{0, 1, 2, 3, 4, 5\}$.

El següent pas per tal de trobar les probabilitats de transició p_{ij} és definir les transicions $i \rightarrow j$ entre estats, on $i, j \in I$, i l'espai de totes les configuracions possibles $\Omega = \{(N_1, N_2, N_3): 0 \leq N_1 \leq N_2 \leq N_3 \leq 5\}$ que corresponen a cadascuna d'aquestes. Així doncs, les transicions amb $j < i$ no tindran configuracions assignades pel fet de que no existeix retrocés entre estats.

Per exemple, les configuracions de les transicions que tenen com estat de sortida el 3 (tenir de 3 a 6 arítmies) són:

Transició	Configuracions
3 a 0	No factible
3 a 1	No factible
3 a 2	No factible
3 a 3	$(3,3,3), (3,3,4), (3,3,5), (0,3,3), (1,3,3), (2,3,3)$
3 a 4	$(3,4,4), (3,4,5), (0,3,4), (1,3,4), (2,3,4), (3,3,4)$
3 a 5	$(3,5,5), (0,3,5), (1,3,5), (2,3,5), (3,3,5)$

A l'Annex 2 podem veure les configuracions corresponents a totes les transicions possibles.

- Observació -

Notem que hi ha configuracions que pertanyen a transicions diferents, com és el cas de la (3,4,4) que apareixerà a la transició $3 \rightarrow 4$ i a la $4 \rightarrow 4$. A continuació es veurà com afecta aquest fet a l'estimació de les probabilitats de transició.

Estimació puntual de les probabilitats de transició

Definim $k_{ij|(N_1, N_2, N_3)}$ com el nombre de pacients que segueixen la configuració (N_1, N_2, N_3) en la transició $i \rightarrow j$. Recordem que les configuracions de cada individu les havíem definit com (n_1, n_2, n_3) , i és en aquestes on ens hem de fixar per fer el recompte.

- Observació -

Per les configuracions de la forma (i, i, i) on $i = 0, 1, 2, 3, 4, 5$ cal agafar el doble del nombre de pacients comptabilitzats ja que es considera que fan dues transicions $i \rightarrow i$, una entre els mesos 1 i 2 i l'altre entre els mesos 2 i 3.

Tot seguit, definim k_{ij} com el nombre de pacients que tenen una configuració que correspon a la transició $i \rightarrow j$. Aquesta variable es calcula fent la suma de les $k_{ij|(N_1, N_2, N_3)}$ per a tot (N_1, N_2, N_3) que pertanyi a la transició $i \rightarrow j$. Així, ja podem estimar les probabilitats de transició a un pas p_{ij} , és a dir, la probabilitat d'estar en un mes determinat a l'estat i i en el següent a l'estat j :

$$\hat{p}_{ij} = \frac{k_{ij}}{\sum_j k_{ij}} \forall i, j = 0, 1, 2, 3, 4, 5, j \geq i.$$

- Observació -

Aquesta fórmula pot ser interpretada com el càlcul de casos favorables dividit casos possibles.

Es pot veure com, amb aquest mètode d'estimació, alguns pacients queden comptabilitzats diversos cops. Aquest fet implica que les estimacions de les probabilitats de transició no siguin independents entre elles.

A continuació es presenta un exemple on es calcula l'estimació puntual de la probabilitat de transició de l'estat 0 al 0:

TRANSICIONS i j	CONFIGURACIONS (N1,N2,N3)	Kij (N1,N2,N3)	Kij	Estima Pij
0 0	(0,0,0)x2	24x2=48	55	55/70=0.786
	(0,0,1)	3		
	(0,0,2)	1		
	(0,0,3)	0		
	(0,0,4)	2		
	(0,0,5)	1		

Taula 1. Exemple de càlcul de l'estimació p_{00}

La transició $0 \rightarrow 0$ correspon a no tenir cap arítmia acumulada al final d'un mes determinat i tampoc acumular-ne cap fins al final del mes següent. Per aquesta transició es defineixen sis configuracions possibles: anar de 0 a 0 entre el primer i el segon mes i en el tercer mes estar a qualsevol estat 0, 1, 2, 3, 4 o 5.

Cal notar que a la configuració (0,0,0) es realitzen dues transicions $0 \rightarrow 0$. A continuació es comptabilitzen el nombre de pacients que tenen cadascuna de les configuracions, per exemple, hi han 3 individus que segueixen la configuració (0,0,1) en el tres primers mesos de l'estudi. Tot seguit es calcula el total de pacients de les configuracions definides per la transició $0 \rightarrow 0$ que en aquest cas són 55. Si repetim aquest mètode per a la resta de transicions, n'obtenim que hi ha un total de 70 pacients comptabilitzats per a alguna de les transicions amb origen 0. Per tant, per a fer l'estimació puntual de la probabilitat de transició $0 \rightarrow 0$ cal que fem el quocient entre els 55 (casos favorables) i el 70 pacients (casos possibles), obtenint que $\hat{p}_{00} = 0.786$.

A l'Annex 2 es troba la taula amb totes les transicions, les configuracions, les $k_{ij(N1,N2,N3)}$, les k_{ij} i les estimacions puntuals de les p_{ij} .

2.5.3. Mètode de Jackknife

El Jackknife és un mètode que ens serveix per obtenir la precisió d'un estimador. Ens proporciona una resposta ràpida a problemes de difícil tractament algebraic tot i que hi afegeix molta feina de càlcul.

Aquest mètode va ser desenvolupat per Quenouille com a procediment per a reduir el biaix d'un estimador i batejat per Tukey com a Jackknife (literalment, navalla de diverses utilitats), generalitzant-lo com a mètode general d'estimació.

Suposem que estimem p_{ij} mitjançant un estimador \hat{p}_{ij} . A continuació, presentem el mètode a seguir:

1. S'elimina de la mostra a un dels 59 pacients, x_s , i es calcula el valor de l'estimador, seguint el mètode exposat anteriorment, amb la mostra de grandària 58 que queda. Anomenarem $X_{(s)}$ a la mostra sense el pacient x_s i $\hat{p}_{ij(s)}$ a l'estimador puntual de la probabilitat de transició de l'estat i a l'estat j , obtingut amb aquesta mostra.
2. Repetim el procediment 59 cops (és la nostra grandària mostral n) per obtenir els estimadors $\hat{p}_{ij(1)}$, $\hat{p}_{ij(2)}$, $\hat{p}_{ij(3)}$, ..., $\hat{p}_{ij(59)}$, obtinguts eliminant el primer, segon, ... pacient i calculant l'estimador amb la mostra que queda.
3. Un cop es tenen les 59 estimacions, basades en 58 observacions, calculem la mitjana i la desviació tipus de les probabilitats estimades segons Jackknife⁽²⁾:

$$\bar{\hat{p}}_{ij} = \frac{1}{n} \sum_{s=1}^n \hat{p}_{ij(s)} \longrightarrow \text{la mitjana de les estimacions puntuals de les probabilitats de transició } i \rightarrow j \text{ es calcula fent el sumatori, per a les } n \text{ mostres, de les estimacions d'aquestes i dividint per } n.$$

$$\hat{\sigma}_{JACK} = \left(\frac{n-1}{n} \sum_{s=1}^n (\hat{p}_{ij(s)} - \bar{\hat{p}}_{ij})^2 \right)^{1/2}$$

→ la desviació tipus es calcula fent l'arrel del sumatori, per a les n mostres, de les diferències entre l'estimació puntual de la probabilitat de transició $i \rightarrow j$ a la mostra s i la mitjana calculada abans, multiplicat pel factor $(n-1)/n$ (enlloc de $1/(n-1)$, com es fa habitualment, ja que és la modificació que fa el mètode de Jackknife).

4. Ara només cal construir l'interval de confiança:

$$p_{ij} \in [\bar{\hat{p}}_{ij} - t_{n-1}^{1-\alpha} \hat{\sigma}_{JACK}; \bar{\hat{p}}_{ij} + t_{n-1}^{1-\alpha} \hat{\sigma}_{JACK}]$$

Així, la probabilitat de transició entre els estats i,j estarà amb una confiança $1-\alpha$ dins l'interval centrat en la mitjana de les estimacions puntuals de les probabilitats p_{ij} .

Així doncs, per tal de construir els intervals de confiança del 95% per a les estimacions de les p_{ij} utilitzarem aquest mètode ja que ens manté l'estructura de correlacions de la matriu de probabilitats de transició.

Per a realitzar-ho, hem creat una Macro per Minitab (veure *Annex 3*) que ens facilita els càlculs: hem creat tres columnes que, per files, constitueixen totes les configuracions possibles (N1,N2,N3), que fan un total de 106 combinacions. Aquestes ens serviran per fer el recompte dels pacients que les segueixen. També es tenen altres tres columnes que, per files, corresponen a les configuracions observades dels 59 individus (N11,N22,N33). Tot seguit s'inicia un bucle que, en cada iteració, ens treurà de la mostra a un dels pacients, x_s , de manera que no el tindrà en compte per trobar els valors de $k_{ij(N1,N2,N3)}$. Així doncs, comencem per la iteració on no es té en compte la configuració del pacient 1, és a dir, no es té en compte la fila 1. Ara cal fer les comparacions entre la resta de files (pacients) per saber quants individus segueixen la configuració c (on $c = 1, \dots, 106$). Aquesta comparació es realitzarà veient si les columnes $N1 = N11$, $N2 = N22$ i $N3 = N33$. Si totes les parelles tenen el mateix

estat, és a dir, es compleix la condició lògica, comptarem aquell pacient en el $k_{ij(N1,N2,N3)}$ corresponent a la configuració c . Un cop haguem fet el recorregut de totes les configuracions possibles i , per tant, tinguem la mostra $X_{(s)}$, passarem a crear la següent mostra on no es tindrà en compte al segon pacient i així successivament fins a tenir les 59 mostres.

Un cop tenim els $k_{ij(N1,N2,N3)}$ per a totes les mostres de grandària 58 calculem els k_{ij} (nombre de pacients que tenen una configuració que pertany a la transició $i \rightarrow j$) i les estimacions de les p_{ij} igual que havíem fet anteriorment per a fer les estimacions puntuals. Aquest procediment s'ha de repetir per a totes les mostres de manera que tindrem 59 estimacions de les p_{ij} . Podem esperar que no hi hagi molta variabilitat en les \hat{p}_{ij} ja que les mostres només es diferencien en un sol individu.

Ara, cal que calculem la mitjana i la desviació tipus de les probabilitats i ja podem trobar els intervals de confiança del 95%.

2.5.4. Intervals de confiança amb la Distribució Binomial

Les estimacions de les probabilitats de transició entre estats es poden calcular també mitjançant la Distribució Binomial, on es comptabilitza el nombre d'aparicions d'un esdeveniment determinat en una seqüència de successos. Així, es considera que el nombre total de successos N_i és el nombre total de pacients que tenen transicions amb origen i (on $i = 0, \dots, 5$), és a dir, seguint la nomenclatura anterior:

$$\sum_j k_{ij} \text{ per } j = 0, \dots, 5.$$

D'altra banda, el nombre d'encerts observats, x_{ij} , correspondran al nombre de pacients que tenen la transició $i \rightarrow j$, és a dir, el nostre k_{ij} .

Si volem estimar el paràmetre p_{ij} , agafem com a estimador puntual la proporció d'èxits en les N_i proves:

$$\hat{p}_{ij} = \frac{x_{ij}}{N_i}$$

Per tal de simplificar, seguirem la següent notació: en N assaigs, el succés E es produeix $x = k$ cops. Els límits de confiança per a $p = x/N$, p_i i p_d que satisfan l'equació:

$$\text{Prob}(p_i < p < p_d | x = k, N) = 1 - 2\alpha$$

són les solucions de:

$$\left\{ \begin{array}{l} \sum_{x=k}^N \binom{N}{x} p_i^x (1-p_i)^{N-x} = \alpha \quad \text{per } p_i \\ \sum_{x=0}^k \binom{N}{x} p_d^x (1-p_d)^{N-x} = \alpha \quad \text{per } p_d \end{array} \right.$$

CAPÍTOL 3

RESULTATS

3.1. NOMBRE D'EPISODIS PER PACIENT

La variable que volem estudiar, nombre d'episodis totals per a cada pacient, no existeix en cap de les dues bases (pacients i episodis), pel què hem hagut de modificar-les.

En primer lloc, hem calculat les freqüències d'arítmies per a cada malalt, a partir de la variable *pacients* de la base de dades *episodis*, creant la variable *numepi* que compta els episodis que ha tingut cada pacient durant el temps que ha estat en seguiment, la qual hem afegit a la base de dades de pacients. Aquesta variable pren valors enters positius, ja que un pacient pot tenir cap arítmia o tenir-ne una o més d'una.

Cal destacar que no tots els pacients han tingut el mateix temps de seguiment pel que hem treballat amb la taxa del nombre d'episodis.

A continuació veurem les corresponents descriptives i l'avaluació de la sobredispersió que observem en aquestes variables.

Tot seguit presentem la descriptiva corresponent a la variable nombre d'episodis:

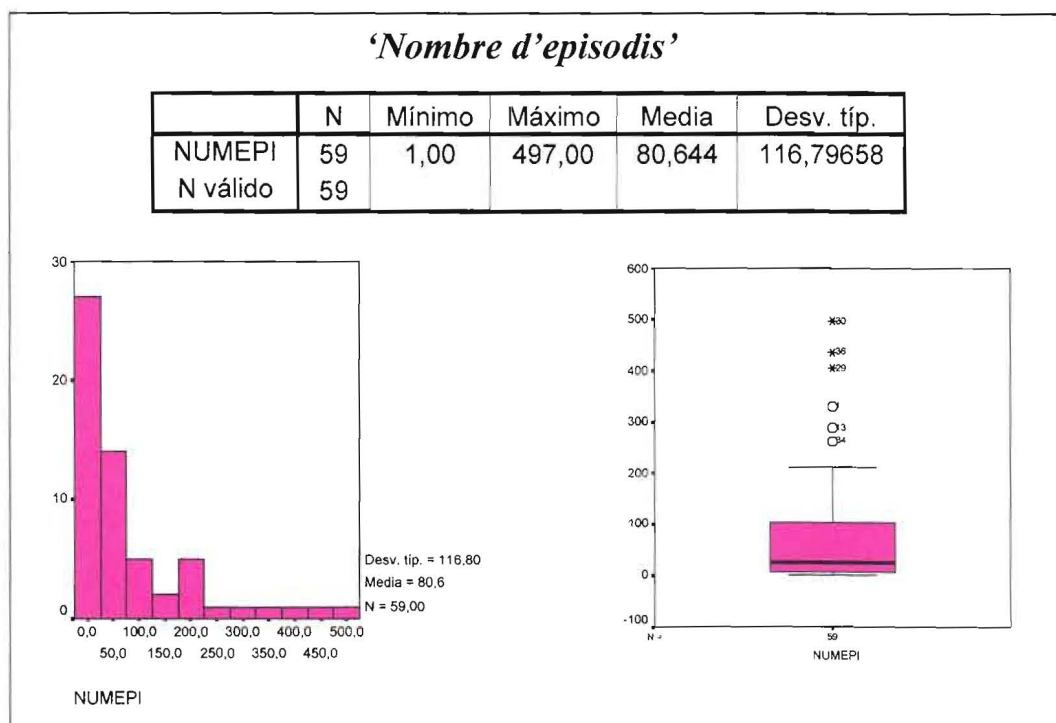


Fig.7. Descriptiva del nombre d'episodis.

Podem observar un decreixement molt important a l'histograma (*figura de l'esquerra*), on la majoria de pacients han tingut menys de 25 arítmies en tot el període de seguiment. Trobem també grups de malalts importants que han patit de 25 a 75, de 75 a 125 i de 175 a 225 episodis. A partir de les 250 arítmies s'observa un únic individu.

Al box-plot (*figura de la dreta*), comprovem la situació anterior. Veiem també que no hi ha simetria i que els quartils 1 i 3 es troben entre 1 i 100 episodis.

Cal destacar que no tots els pacients han tingut el mateix temps de seguiment, el que aporta una sobredispersió important a les dades. Verifiquem doncs, als resultats anteriors, que la variable nombre d'episodis està molt sobredispersionada, ja que la variança és molt més gran que la mitjana: $116,797^2 \gg 80,644$.

$$\text{Índex de dispersió} = \frac{\sigma_{\text{numepi}}^2}{\text{numepi}} = 116,797^2 / 80,644 = 169,158 \gg 1$$

La dificultat afegida de tenir temps de seguiment diferents per pacient, l'hem solucionat treballant amb la taxa del nombre d'episodis. Hem dividit el nombre d'episodis pel nombre de mesos de seguiment de cada pacient i hem creat la variable taxa del nombre d'episodis *taxaepi*.

L'avantatge principal de treballar amb la taxa és que eliminem el fet de què hi hagi pacients amb més arítmies que altres malalts degut a què els seus temps de seguiment hagin estat majors i donin lloc a aquest fet. Aquesta variable pren valors reals positius ja que totes les persones han patit arítmies.

Obtenim els següents resultats:

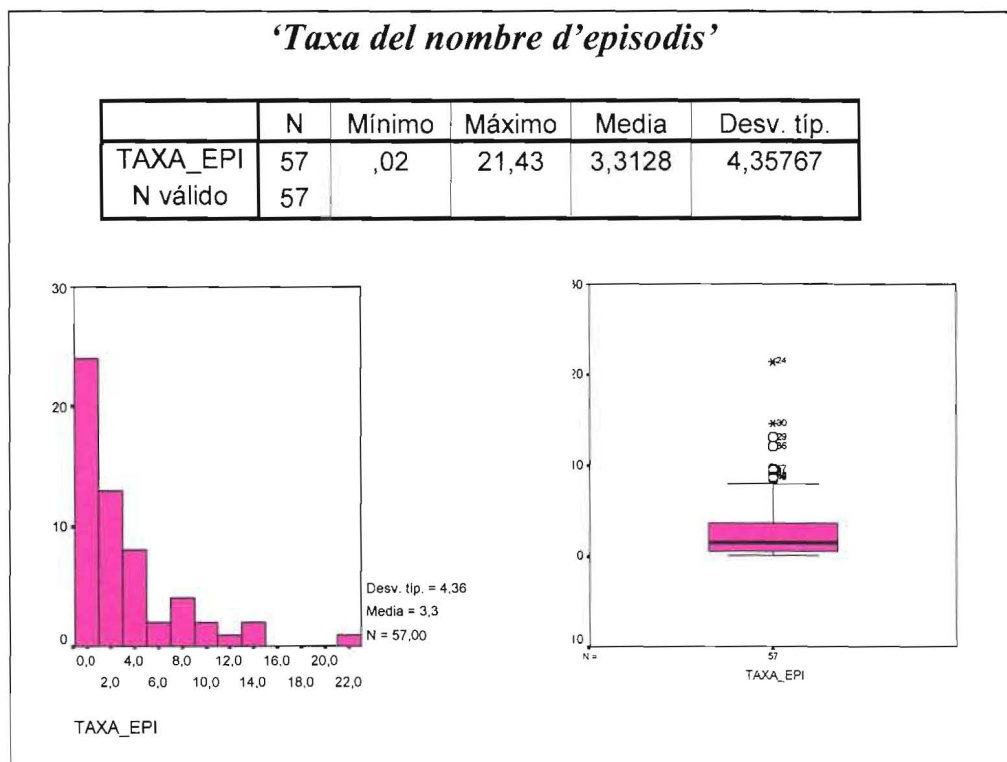


Fig.8. Descriptiva de la taxa del nombre d'episodis.

Podem observar un decreixement més lleuger que abans a l'histograma (*figura de l'esquerra*), la majoria dels pacients estan al voltant d'una taxa igual a zero.

Al box-plot (*figura de la dreta*) veiem, com abans, que no hi ha simetria i que la dispersió es troba por sobre de la mediana.

Tot així continuem observant sobredispersió a la variable:

$$\text{Índex de dispersió} = \frac{\sigma_{taxae\pi i}^2}{taxae\pi i} = 4'358^2/3'313 = 5'733 \gg 1$$

Aquesta sobredispersió pot venir provocada per una heterogeneïtat de la població i, per tant, podríem pensar que la taxa del nombre d'episodis fos una mixtura de Poissons. Els passos següents d'aquest anàlisi es discuteixen al *punt 4.3*.

3.2. EVOLUCIÓ DEL TEMPS ENTRE ARÍTMIES

Estudiarem la relació entre dos arítmies consecutives: temps que passa des de què es produeix una arítmia fins que apareix la següent.

Com hem observat anteriorment, tots els pacients van tenir com a mínim una arítmia en tot el període de tractament. Tot i així, per a calcular els temps t_{ij} hem tingut en compte 47 pacients, ja que pels 12 malalts restants sabem que hi ha 11 que han tingut només una arítmia, fet que no ens aporta cap informació dels temps entre la primera i segona i entre la segona i la tercera arítmia. Del pacient restant, tenim constància de que ha tingut 3 arítmies però només tenim registrat el temps fins al primer episodi.

Tot seguit presentarem l'estudi on analitzem la variable en qüestió i dues transformacions: sumes, diferències i els seus logaritmes, per tal de trobar la millor manera d'explicar l'evolució del temps.

- Observació -

Cal destacar que a l'hora d'estudiar els logaritmes de les diferències, hem hagut de sumar una unitat a la variable corresponent, ja que moltes vegades ens donava zero.

3.2.1. Temps entre arítmies

A continuació, es presenten els gràfics on es relacionen els temps entre arítmies:

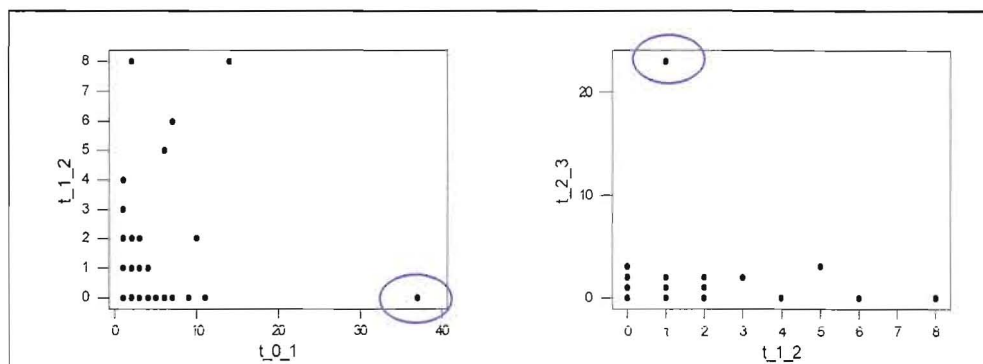


Fig. 9. Plot dels temps entre arítmies.

Com veiem a la figura, hi ha una observació que distorsiona els gràfics. Per tal de tenir una millor visualització de què és el que passa, hem tret l'observació 20 ja que el temps fins a la primera arítmia és molt gran ($t_{01} = 37$) i a més la resta són zero ($t_{12} = 0$ i $t_{23} = 0$), el que vol dir que l'individu que correspon a aquesta observació ha patit les 3 arítmies en al mateix mes. Cal comentar que hem tret l'observació per la construcció dels gràfics, però l'hem inclòs en els càlculs numèrics.

Gràfics on es relacionen els temps entre arítmies, sense l'observació 20:

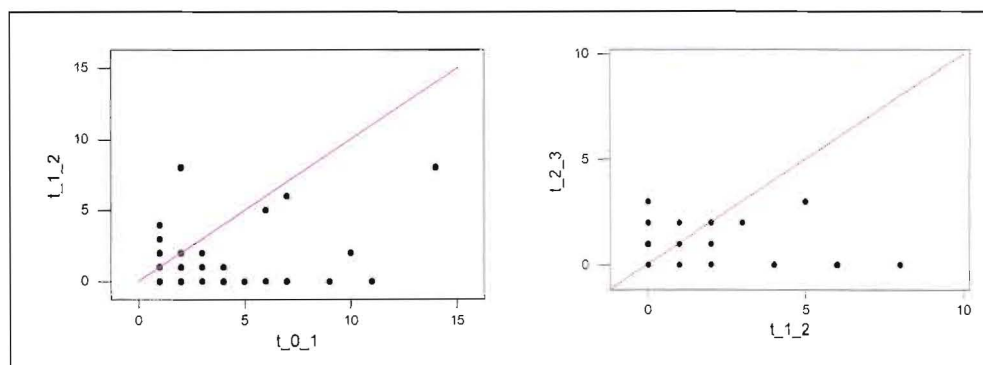


Fig. 10. Plot dels temps entre arítmies.

Observem així que, a mesura que el temps entre les arítmies és major, la dispersió augmenta. La majoria d'observacions es troben en espais de temps curts entre arítmies.

Per facilitar la interpretació presentem la descriptiva de les variables:

Variable	N	N*	Mean	Median	TrMean	StDev
t_0_1	47	12	3,809	2,000	2,930	5,785
t_1_2	47	12	1,085	0,000	0,814	2,009
t_2_3	47	12	0,979	0,000	0,465	3,391

Variable	SE Mean	Minimum	Maximum	Q1	Q3
t_0_1	0,844	1,000	37,000	1,000	4,000
t_1_2	0,293	0,000	8,000	0,000	1,000
t_2_3	0,495	0,000	23,000	0,000	1,000

Fig.11. Descriptiva dels temps entre arítmies.

Observant les mitjanes de les tres variables, veiem que el temps entre dos arítmies va disminuint a mida que hi apareixen més arítmies. És a dir, el temps fins a la primera és major que el temps entre la primera i la segona, i aquest és major que el temps entre la segona i la tercera arítmia ($t_{01} > t_{12} > t_{23}$).

Per a comprovar que els temps es redueixen, hem realitzat la comparació de mitjanes d'aquestes variables mitjançant l'anàlisi de la variança per a dades aparellades per pacient:

$$\begin{cases} H_0 : \mu_{t_{01}} = \mu_{t_{12}} = \mu_{t_{23}} \\ H_1 : \exists i, j \text{ tal que } \mu_i \neq \mu_j \end{cases}$$

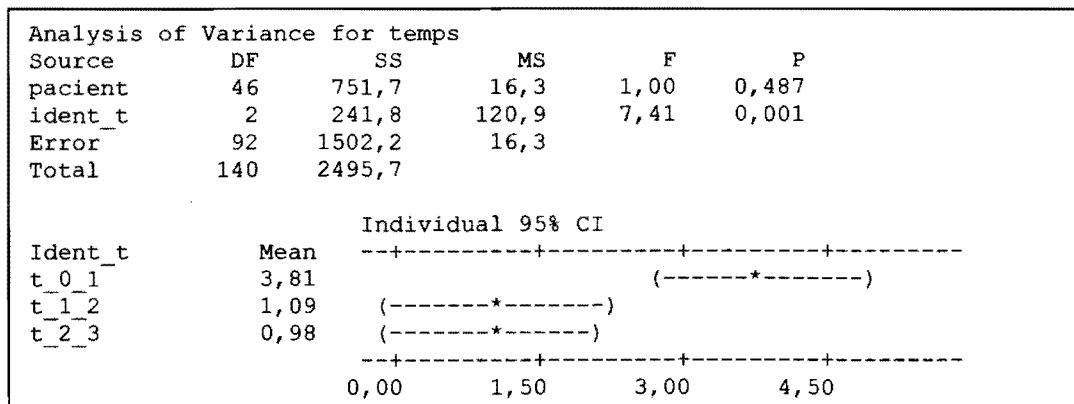


Fig.12. Test de comparació de les mitjanes dels temps entre arítmies.

Veiem que hi ha diferències significatives entre els temps entre arítmies (t_{01} , t_{12} i t_{23}), ja que tenim un p-valor de 0'001, amb el que rebutgem H_0 , d'igualtat entre mitjanes.

3.2.2. Estudi de l'additivitat del temps entre arítmies

A continuació presentem l'estudi de les sumes i les diferències del temps entre arítmies, ja que les interpretacions ens seran més interessants. En els gràfics de Bland-Altman, les abscisses són les sumes entre el temps d'un episodi i el següent ($t_{ij} + t_{i+1,j+1}$) i a les ordenades es troben les seves diferències ($t_{i+1,j+1} - t_{ij}$). Quan els temps entre dues arítmies consecutives és el mateix ($t_{ij} = t_{i+1,j+1}$) queda determinada la recta d'equivalència.

Ara bé, la interpretació en aquests tipus de gràfics és més clara, ja que amb les sumes es pot veure la magnitud del temps i amb les diferències es pot comprovar la discrepància entre els temps t_{ij} i $t_{i+1,j+1}$ per a una mateixa observació.

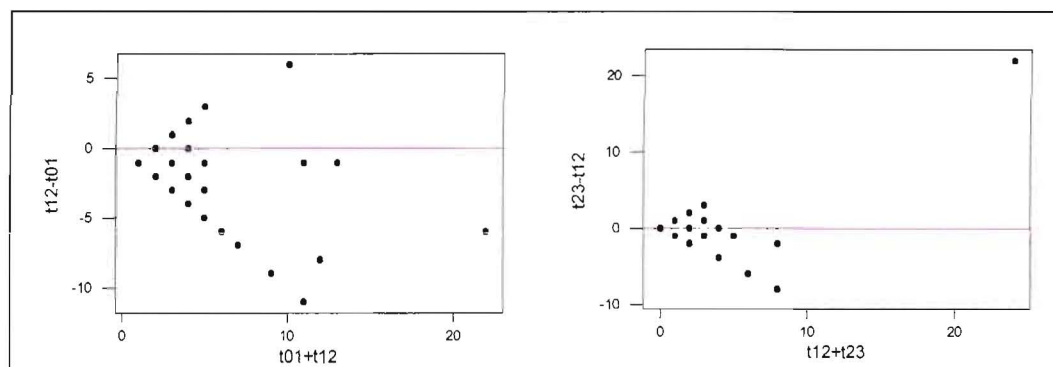


Fig. 13. Plot de les sumes i les diferències dels temps entre arítmies.

Observant el gràfic de l'esquerra, veiem que la majoria de punts queden per sota de la recta d'equivalència, és a dir, que el temps fins a la primera arítmia és major que el temps entre la primera i la segona ($t_{12} < t_{01}$). Amb això podem dir que els temps es redueixen per a gairebé totes les observacions. Al gràfic de la dreta no es veu de forma tan clara la situació anterior.

Presentem la descriptiva de les diferències:

Variable	N	N*	Mean	Median	TrMean	StDev
t12-t01	47	12	-2,723	-1,000	-2,070	5,900
t23-t12	47	12	-0,106	0,000	-0,326	3,891

Variable	SE Mean	Minimum	Maximum	Q1	Q3
t12-t01	0,861	-37,000	6,000	-3,000	-1,000
t23-t12	0,568	-8,000	22,000	0,000	0,000

Fig. 14. Descriptiva de les diferències dels temps entre arítmies.

Veiem que la diferència entre el temps que hi ha entre tenir la primera arítmia i el temps entre la primera i la segona ($t_{12} - t_{01} = -2,723$) és més gran, en valor absolut, que la diferència entre el temps que hi ha entre tenir la primera i la segona i entre la segona i la tercera ($t_{23} - t_{12} = -0,106$).

Podem observar en ambdues variables ($t_{12}-t_{01}$ i $t_{23}-t_{12}$), que la variança de la diferència dels temps entre arítmies és molt més gran que la seva mitjana (en valor absolut), prenent valors de $|-2'723| \ll 5'90^2 = 34'81$ i $|-0'106| \ll 3'891^2 = 15'139$ respectivament. A més, havíem vist en els gràfics anteriors que la dispersió augmentava amb el valor mig. Per aquesta raó, veurem més endavant l'anàlisi de les variables prenent logaritmes.

A continuació realitzarem un test per tal de veure si la mitjana de les diferències entre els temps entre arítmies és zero o no:

$$\begin{cases} H_0 : \mu_{t_{21}-t_{01}} = 0 \\ H_1 : \mu_{t_{21}-t_{01}} \neq 0 \end{cases} \quad \begin{cases} H_0 : \mu_{t_{23}-t_{12}} = 0 \\ H_1 : \mu_{t_{23}-t_{12}} \neq 0 \end{cases}$$

Test of mu = 0 vs mu not = 0				
Variable	N	Mean	StDev	SE Mean
t12-t01	47	-2,723	5,900	0,861
t23-t12	47	-0,106	3,891	0,568

Variable	95,0% CI	T	P
t12-t01	(-4,456; -0,991)	-3,16	0,003
t23-t12	(-1,249; 1,036)	-0,19	0,852

Fig. 15. Prova d'hipòtesi d'una mitjana: $\mu=\mu_0$ per les diferències dels temps entre arítmies.

Com es pot veure a la *figura 15*, podem rebutjar que la mitjana de les diferències del temps entre tenir la primera i la segona arítmia i el temps que passa fins a tenir la primera ($\mu_{t_{21}-t_{01}}$) sigui zero, ja que tenim un p-valor de 0'003. En canvi, no hi han evidències per a rebutjar que la mitjana de les diferències del temps entre la segona i la tercera arítmia i entre la primera i la segona ($\mu_{t_{23}-t_{12}}$) sigui zero, ja que tenim un p-valor de 0'852.

Així, es pot veure que, la diferència entre el temps que passa entre la primera i la segona arítmia i el temps que passa fins a la primera ($t_{12} - t_{01}$) té una mitjana de -2'723 (-4'456 ; -0'991), negativa i significativament diferent de zero. Per tant, el segon episodi triga en aparèixer entre 1 i 4'5 mesos menys que el primer, amb el que podem comprovar que el temps es redueix.

En canvi, la diferència entre el temps que passa entre la segona i la tercera arítmia i el temps que passa entre la primera i la segona ($t_{23} - t_{12}$) té una mitjana de -0'106 (-1'249 ; 1'036). Veiem que el tercer episodi triga en aparèixer entre 1'2 mesos menys i 1 més que el segon, l'interval inclou el zero. Així doncs, no observem cap reducció de temps.

A continuació presentem el test de normalitat per tal de comprovar que la diferència de t_{12} i t_{01} compleix amb les premisses de normalitat:

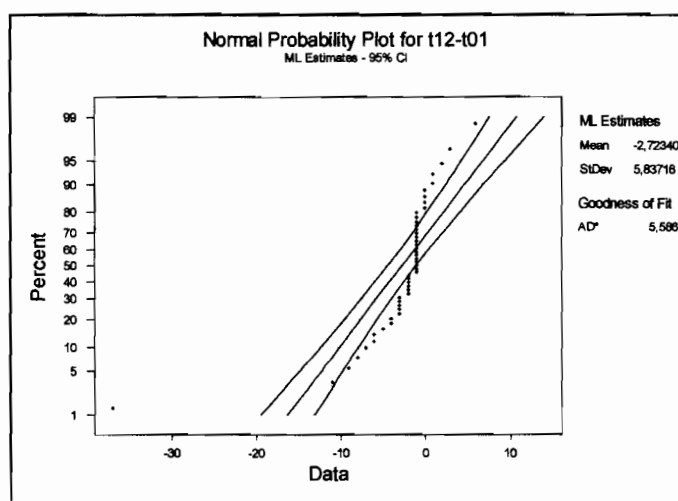


Fig. 16. Gràfic de normalitat de la variable $t_{12}-t_{01}$

Com es pot veure al gràfic de normalitat del estudi de l'additivitat del temps entre arítmies, la diferència del temps t_{12} i t_{01} no segueix una normal.

3.2.3. Estudi de l'additivitat en escala logarítmica del temps entre arítmies

Com hem comentat a l'anàlisi anterior, la variança de la diferència dels temps entre arítmies és molt més gran que la seva mitjana i la dispersió augmenta amb el valor mig. Per un altre banda, hem vist que el gràfic de normalitat de les sumes i les diferències dels temps entre arítmies no s'ajusta bé a una normal. A continuació realitzem transformacions logarítmiques per tal de resoldre aquestes qüestions.

Presentem l'estudi fent les sumes i les diferències dels logaritmes dels temps entre arítmies:

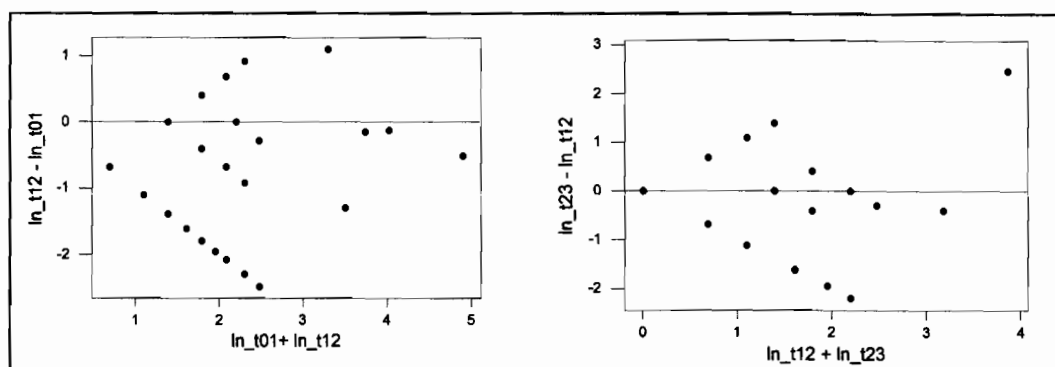


Fig. 17. Plot de les sumes i les diferències dels logaritmes dels temps entre arítmies.

Es repeteix la situació anterior, la majoria de punts queden per sota de la recta d'equivalència ($t_{12} < t_{01}$) al gràfic de l'esquerra. Al gràfic de la dreta no es veu tan clara aquesta situació. Igualment en ambdós casos podem dir que, a mida que el temps es va fent més gran la divergència augmenta: heterocedasticitat.

Descriptiva dels diferències dels logaritmes dels temps entre arítmies:

Variable	N	N*	Mean	Median	TrMean	StDev
ln_t12 - ln_t01	47	12	-0,807	-0,690	-0,787	0,887
ln_t23 - ln_t12	47	12	-0,111	0,000	-0,109	0,802

Variable	SE Mean	Minimum	Maximum	Q1	Q3
ln_t12 - ln_t01	0,129	-3,640	1,100	-1,390	-0,290
ln_t23 - ln_t12	0,117	-2,200	2,480	0,000	0,000

Fig. 18. Descriptiva de les diferències dels logaritmes dels temps entre arítmies.

Veiem que la diferència entre els logaritmes dels temps que hi ha entre tenir la primera arítmia i entre la primera i la segona ($\ln t_{12} - \ln t_{01} = -0,807$) és més gran, en valor absolut, que la diferència entre els logaritmes dels temps que hi ha entre la primera i la segona i entre la segona i la tercera ($\ln t_{23} - \ln t_{12} = -0,111$).

Pel que fa a la primera variable comentada ($\ln t_{12} - \ln t_{01}$), la varianza és més petita que la mitjana (en valor absolut), prenent els següents valors $0'887^2 = 0'787 < |-0,807|$ respectivament. Pel que fa a la segona variable ($\ln t_{23} - \ln t_{12}$), seguim amb la situació de l'estudi sense logaritmes on la varianza és més gran que la mitjana ($0'802^2 = 0'643 > |-0'111|$).

Realitzem un test per veure si les diferències dels logaritmes dels temps són significativament diferents de zero o no:

$$\begin{cases} H_0 : \mu_{\ln t_{21} - \ln t_{01}} = 0 \\ H_1 : \mu_{\ln t_{21} - \ln t_{01}} \neq 0 \end{cases} \quad \begin{cases} H_0 : \mu_{\ln t_{23} - \ln t_{12}} = 0 \\ H_1 : \mu_{\ln t_{23} - \ln t_{12}} \neq 0 \end{cases}$$

Test of mu = 0 vs mu not = 0				
Variable	N	Mean	StDev	SE Mean
ln_t12 - ln_t01	47	-0,807	0,887	0,129
ln_t23 - ln_t12	47	-0,111	0,802	0,117

Variable	95,0% CI	T	P
ln_t12 - ln_t01	(-1,068; -0,547)	-6,24	0,000
ln_t23 - ln_t12	(-0,346; 0,125)	-0,95	0,349

Fig. 19. Prova d'hipòtesi d'una mitjana: $\mu = \mu_0$ per les diferències dels logaritmes dels temps entre aritmies.

Observem que la primera mitjana de la diferència de logaritmes ($\mu_{\ln t_{21} - \ln t_{01}}$) és $-0'807$, negativa i significativament diferent de zero, ja que el p-valor és zero (podem rebutjar la hipòtesi nul·la). En els intervals de confiança, veiem que el segon episodi triga en aparèixer entre 34'369% ($e^{-1'068}$) i un 57'868% ($e^{-0'547}$) menys que el primer. Amb el que podem comprovar una reducció del temps.

En canvi, la segona mitjana ($\mu_{\ln t_{23} - \ln t_{12}}$) és $-0'95$ ($-0'346$; $0'125$) no significativament diferent de zero, ja que el p-valor és 0'349 (no tenim evidències per rebutjar H_0). A més, el seu interval de confiança inclou el zero. Per tant, podem dir que no hi ha reducció del temps a partir de la segona arítmia.

Treballant amb logaritmes podem comprovar que:

- el temps entre arítmies es redueix després del primer episodi:

$$\ln t_{12} - \ln t_{01} = \ln(t_{12} / t_{01}) < 0 \Rightarrow t_{12} / t_{01} < 1 \Rightarrow t_{12} < t_{01}$$

- a partir de la segona, el temps entre arítmies s'estabilitza:

$$\mu_{\ln t_{23} - \ln t_{12}} \in [-0'346; 0'125]$$

$$\text{Si } \ln t_{23} - \ln t_{12} = \ln(t_{23} / t_{12}) = -0'346 \Rightarrow t_{23} / t_{12} = e^{-0'346} = 0'7075$$

$$\text{Si } \ln t_{23} - \ln t_{12} = \ln(t_{23} / t_{12}) = 0'125 \Rightarrow t_{23} / t_{12} = e^{0'125} = 1'133$$

$\mu_{t_{23} / t_{12}} \in [0'7075; 1'133] \Rightarrow$ L'interval inclou l' 1, fet que contempla la possibilitat de que $t_{23} \approx t_{01}$.

A continuació presentem el test de normalitat de per tal de comprovar que la diferència entre t_{12} i t_{01} compleix amb les premisses de normalitat:

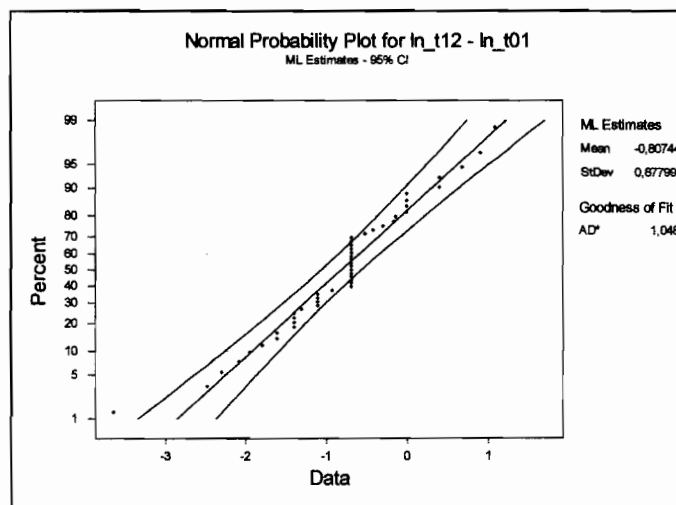


Fig.20. Gràfic de normalitat de la variable $\ln t_{12} - \ln t_{01}$.

Com es pot veure, tant a l'estudi de l'additivitat del temps entre arítmies com en l'estudi de l'additivitat en escala logarítmica, arribem a les mateixes conclusions: el temps lliure d'arítmies disminueix.

Ara bé, en el segon estudi, el gràfic de normalitat surt molt millor, pel què triaríem aquest model per explicar l'evolució del temps entre arítmies.

- Observació -

Cal fer notar que el fet de que t_{01} sigui censurada per l'esquerra, reforça els nostres resultats, ja que el temps que hi ha entre la última arítmia abans de l'implant i aquest (del que no tenim constància a la base de dades) serà més gran que el temps fins a la primera arítmia després de l'implant (t_{01}).

- Observació -

Com hem dit al començament de l'estudi, el nombre d'observacions a la base de dades es redueix de 59 a 47.

Una possibilitat d'avaluar si els temps que falten d'aquests 12 malalts pot invertir la conclusió a la que hem arribat, és veure quin valor mitjà del temps que passa entre la primera arítmia i la segona (t_{12}) d'aquests malalts faria que els temps que passa entre tenir la primera arítmia i entre la primera i la segona (t_{01} i t_{12}) es facin iguals.

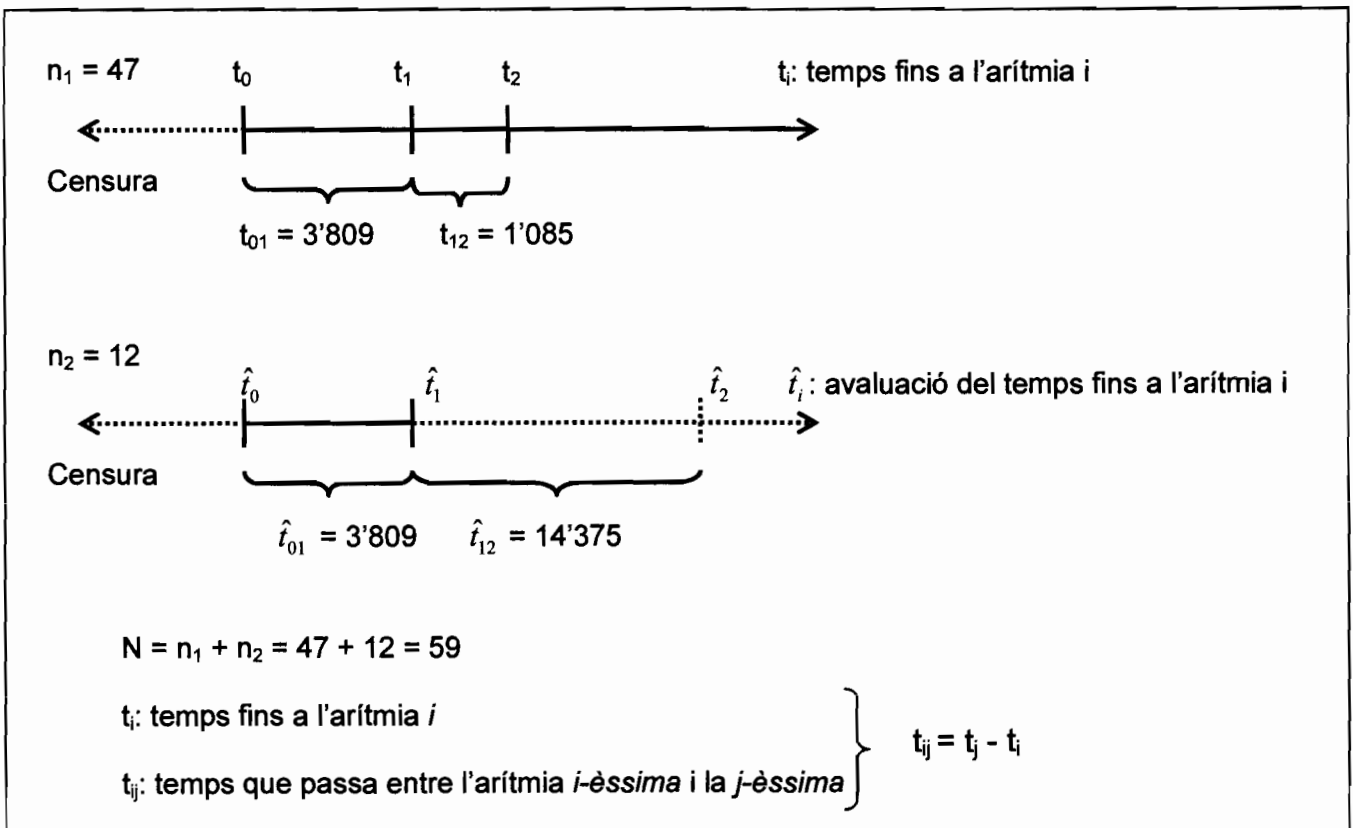


Fig.21. Avaluació del temps.

Avaluació de \hat{t}_{01} com la mitjana dels temps t_{01} : $\hat{t}_{01} = 3'809$

Avaluació de \hat{t}_{12} :

$$\hat{t}_{01} = [(\bar{t}_{12} * n_1) + (\hat{t}_{12} * n_2)] / N$$

$$3'809 = [(1'085 * 47) + (\hat{t}_{12} * 12)] / 59$$

$$\Rightarrow \hat{t}_{12} = [(3'809 * 59) - (1'085 * 47)] / 12 = 14'478$$

[NOTA: respecte al pacient que ha tingut els tres episodis però no tenim el temps de seguiment, com que ha tingut les tres arítmies dins del temps de seguiment sabem que el temps entre arítmies no serà més gran que el que hem avaluat.]

- Observació -

A l'Annex 4 presentem un estudi en el què hem suposat que els 12 pacients, dels que hi ha manca d'informació. Així, hem assignat el valors a les variables t_{ij} , tal que t_{01} serà el temps de seguiment, ja que aquest és el temps mínim en el que es produirà una arítmia i $t_{12} = t_{23} = 0$.

3.3. MODELITZACIÓ DEL NOMBRE D'ARÍTMIES

A continuació es presenten els resultats que hem obtingut construint la cadena de Markov que ens modela el nombre d'arítmies per pacient. Per fer-ho hem seguit la metodologia exposada al Capítol 2.

3.3.1. Matriu P de probabilitats de transició

La matriu de probabilitats de transició P està formada per les probabilitats estimades \hat{p}_{ij} . Aquesta matriu és triangular (té zeros per sota de la diagonal) ja que aquestes posicions corresponen a les p_{ij} amb $j < i$. També es pot observar que es compleix la propietat de les cadenes de Markov de què els sumatoris per files són 1:

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} 0'786 & 0'100 & 0'014 & 0'043 & 0'043 & 0'014 \\ & 0'600 & 0'266 & 0'000 & 0'067 & 0'067 \\ & & 0'330 & 0'670 & 0'000 & 0'000 \\ & & & 0'500 & 0'375 & 0'125 \\ & & & & 0'460 & 0'540 \\ & & & & & 1'000 \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

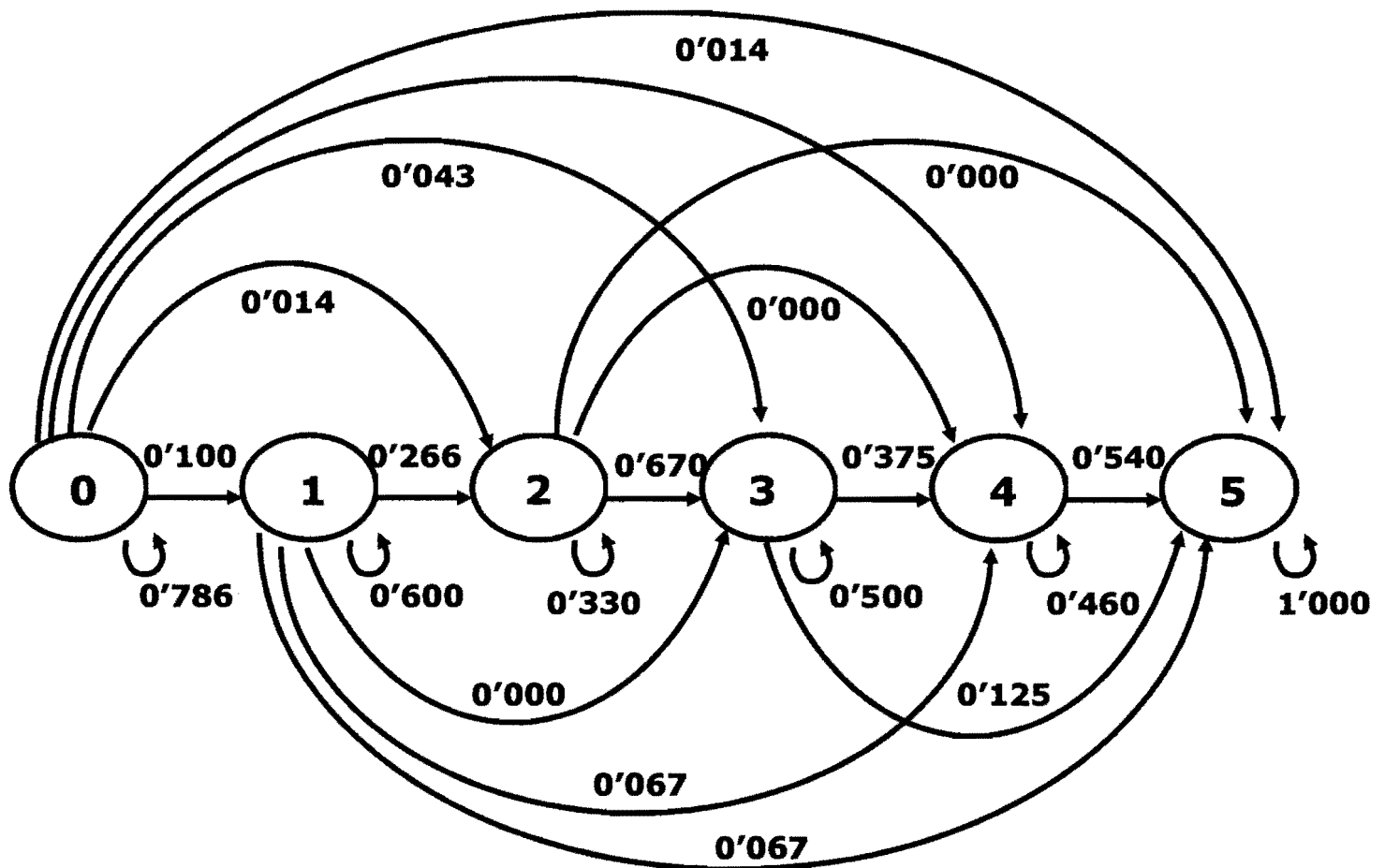
A continuació, s'interpreten una a una les estimacions puntuals d'aquestes probabilitats. L'ordre d'exposició segueix la diagonal.

- L'estimació puntual de la probabilitat de no tenir cap arítmia en el mes següent si no s'ha tingut cap és del 78'6%. Per tant, la de tenir alguna arítmia en el mes següent si no s'ha tingut cap és del 21'4%.
- La de no tenir cap si s'ha tingut una en el mes anterior és del 60%. Per tant, la de tenir-ne alguna és del 40%.
- La de no tenir cap si s'han tingut dues en el mes anterior és del 33%. Per tant, la de tenir-ne alguna és del 67%. Com que l'estimació puntual de passar de dues arítmies a de "set a vint" i de dues a "més de vint" és zero, podem concretar més i dir que aquest 67% anterior correspon a tenir de "tres a sis" arítmies acumulades en el mes següent si en l'anterior s'han tingut dues.
- L'estimació puntual de la probabilitat d'acumular més de sis arítmies en el mes següent si actualment s'han tingut de "tres a sis" és del 50%. La de continuar

tenint de "tres a sis" arítmies quan s'han tingut també de tres a sis en el mes anterior és del 50%.

- La d'acumular "més de vint" arítmies en el mes següent si s'han tingut de "set a vint" és del 54%. I la de continuar tenint de "set a vint" arítmies quan s'han tingut de "set a vint" en el mes anterior és del 46%.
- La de tenir una arítmia en el mes següent quan no s'ha tingut cap en el mes anterior és del 10% i quan s'ha tingut una és del 26'6%.
- L'estimació puntual de la probabilitat de tenir dues arítmies en un mes si en l'anterior no s'ha tingut cap és del 1'4%, de tenir-ne de "tres a sis" és de 4'3%, de tenir-ne de "set a vint" és també del 4'3% i la de tenir "més de vint" arítmies és del 1'4%.
- Si en un mes s'ha tingut una arítmia, l'estimació puntual de la probabilitat de passar a tenir de "tres a sis" arítmies acumulades en el mes següent és del 0%, la de passar a tenir de "set a vint" és del 6'7% i la de tenir "més de vint" és del 6'7%.
- La de passar de tenir de "set a vint" arítmies acumulades en un mes si en l'anterior s'havien acumulat de "tres a sis" és del 37'5% i la d'acumular més de vint és del 12'5%.
- Per últim, cal comentar que si en un mes s'han acumulat "més de vint" arítmies en el mes següent es continuarà tenint més de vint amb una probabilitat del 100%, ja que, per construcció, aquest és un estat absorbent i quan un pacient arriba a ell, aquest hi romandrà al llarg del temps.

3.3.2. Diagrama de la cadena de Markov



Recordem que:

- *estat 0*, correspon a tenir zero arítmies acumulades
- *estat 1*, una arítmia acumulada
- *estat 2*, dues arítmies acumulades
- *estat 3*, de tres a sis arítmies acumulades
- *estat 4*, de set a vint arítmies acumulades
- *estat 5*, correspon a tenir més de vint arítmies.

- Observació -

Cal comentar que les estimacions de les probabilitats que són zero, no caldria posar-les en el diagrama, però com que són estimacions puntuals les deixem.

3.3.3. Estimacions per interval de les probabilitats de transició

Per tal de construir els intervals de confiança del 95% per a les estimacions de les p_{ij} utilitzarem el mètode de Jackknife. El procediment seguit és l'explicat al Capítol 2 de material i mètodes.

La *taula 2* conté les mitjanes de les estimacions de probabilitats de transició, la desviació tipus calculada segons l'estimació del mètode de Jackknife i l'interval de confiança del 95% corresponent a cadascuna d'elles. Cal comentar que només hi apareixen les transicions factibles, és a dir, les que compleixen $j > i$.

Transició i, j	\bar{p}_{ij}	$\hat{\sigma}_{JACK}$	$IC_{95\%} p_{ij}$	
0,0	0,786	0,049	0,687	0,885
0,1	0,100	0,036	0,027	0,173
0,2	0,014	0,014	-0,014	0,043
0,3	0,043	0,025	-0,007	0,093
0,4	0,043	0,024	-0,006	0,092
0,5	0,014	0,014	-0,014	0,043
1,1	0,600	0,132	0,335	0,865
1,2	0,267	0,125	0,016	0,518
1,3	0,000	0,000	*	*
1,4	0,067	0,070	-0,074	0,207
1,5	0,067	0,070	-0,073	0,206
2,2	0,333	0,405	-0,477	1,144
2,3	0,667	0,405	-0,144	1,477
2,4	0,000	0,000	*	*
2,5	0,000	0,000	*	*
3,3	0,500	0,240	0,020	0,979
3,4	0,375	0,218	-0,060	0,811
3,5	0,125	0,137	-0,148	0,398
4,4	0,461	0,172	0,117	0,806
4,5	0,539	0,172	0,194	0,883
5,5	1'00	Absorbent	*	*

Taula 2. Estimacions de les probabilitats per interval.

La matriu P amb les estimacions per interval, tenint present la propietat de les probabilitats que ens limita el seu rang de valors entre 0 i 1, queda de la següent manera:

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} (0'687;0'885) & (0'027;0'173) & (0'000;0'043) & (0'000;0'093) & (0'000;0'092) & (0'000;0'043) \\ & (0'335;0'865) & (0'016;0'518) & (*;*) & (0'000;0'207) & (0'000;0'206) \\ & & (0'000;1'000) & (0'000;1'000) & (*;*) & (*;*) \\ & & & (0'020;0'979) & (0'000;0'811) & (0'000;0'398) \\ & & & & (0'117;0'806) & (0'194;0'883) \\ & & & & & (*;*) \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

3.3.4. Estimacions de les probabilitats amb la Binomial

Amb ajut de les taules dels límits de confiança per la p de la distribució Binomial⁽¹⁰⁾, podem trobar les estimacions puntuals i per interval de les p_{ij} . En primer lloc, agafarem la taula amb la N_i que ens interessa i tot seguit anirem agafant les estimacions que corresponguin amb cadascuna de les x_{ij} . Així obtenim la taula següent (taula 3) on només apareixen les probabilitats de les transicions factibles, és a dir, les que corresponen a $j > i$:

Transició i, j	\hat{p}_{ij} (en %)	$IC_{95\%}$	p_{ij} (en %)
0,0	0'7857	0'6713	0'7848
0,1	0'1000	0'0412	0'1952
0,2	0'0143	0'0004	0'0770
0,3	0'0429	0'0089	0'1202
0,4	0'0429	0'0089	0'1202
0,5	0'0143	0'0004	0'0770
1,1	0'6000	0'3229	0'8366
1,2	0'2667	0'779	0'5510
1,3	0'0000	0'0000	0'2316
1,4	0'0714	0'0018	0'3387
1,5	0'0714	0'0018	0'3387
2,2	0'3333	0'0084	0'9057
2,3	0'6667	0'0943	0'9916
2,4	0'0000	0'0000	0'7076
2,5	0'0000	0'0000	0'7076
3,3	0'5000	0'1570	0'8430
3,4	0'3750	0'0852	0'7551
3,5	0'1250	0'0032	0'5265
4,4	0'4615	0'1922	0'7487
4,5	0'5385	0'2513	0'8087
5,5	1'0000	0'6637	1'0000

Taula 3. Estimació puntual i per interval de les probabilitats de transició segons la Distribució Binomial.

Com es pot observar, les estimacions trobades amb els dos sistemes són molt semblants. Així, per les estimacions amb * posarem les estimacions per interval trobades amb el mètode de la Binomial ja que el Jackknife no ens reflexa la variabilitat existent per manca d'observacions. Cal comentar, però, que aquest càlcul no és correcte ja que la Binomial requereix independència entre les observacions i nosaltres no la tenim ja que hi ha pacients que estan comptabilitzant varies vegades.

La matriu de probabilitats de transició queda de la següent manera:

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} (0'687;0'885) & (0'027;0'173) & (0'000;0'043) & (0'000;0'093) & (0'000;0'092) & (0'000;0'043) \\ & (0'335;0'865) & (0'016;0'518) & (0'000;0'232) & (0'000;0'207) & (0'000;0'206) \\ & & (0'000;1'000) & (0'000;1'000) & (0'000;0'708) & (0'000;0'708) \\ & & & (0'020;0'979) & (0'000;0'811) & (0'000;0'398) \\ & & & & (0'117;0'806) & (0'194;0'883) \\ & & & & & (0'664;1'000) \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

3.3.5. Probabilitats en m passos $P^{(m)}$

Fins ara, hem estudiat les probabilitats de passar d'un estat a un altre en 1 mes. Ara bé, com estaran aquestes probabilitats passat un cert temps? Ens cal estudiar les probabilitats en m passos per tal de veure com evolucionen el nombre d'arítmies que pateixen els pacients en el futur. Tot i així, com que la cadena construïda té l'estat 5 com a estat absorbent, sabem que passat un nombre finit de mesos tots els pacients estaran en aquest (tindran més de vint arítmies) amb probabilitat 1. S'ha decidit que els punts més significatius d'observació són 1 any i 3 anys, és a dir, als 12 i als 36 mesos.

Per trobar les estimacions puntuals d'aquestes probabilitats farem la potència de P a la 12 i a la 36. Així doncs, obtindrem $P^{(12)} = P^{12}$ i $P^{(36)} = P^{36}$:

$$P^{(12)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} 0'0560 & 0'0287 & 0'1800 & 0'0463 & 0'0588 & 0'7926 \\ & 0'0022 & 0'0021 & 0'0118 & 0'0237 & 0'9602 \\ & & 0'0000 & 0'0010 & 0'0047 & 0'9983 \\ & & & 0'0002 & 0'0014 & 0'9983 \\ & & & & 0'0001 & 0'9999 \\ & & & & & 1'0000 \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

$$P^{(36)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} 0'0002 & 0'0001 & 0'0001 & 0'0002 & 0'0002 & 0'9993 \\ & 0'0000 & 0'0000 & 0'0000 & 0'0000 & 1'0000 \\ & & 0'0000 & 0'0000 & 0'0000 & 1'0000 \\ & & & 0'0000 & 0'0000 & 1'0000 \\ & & & & 0'0000 & 1'0000 \\ & & & & & 1'0000 \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

Algunes de les interpretacions que en podem treure d'aquestes matrius són que un pacient que no ha acumulat cap arítmia, després d'un any haurà acumulat més de vint amb una probabilitat del 79'26%, i després de tres anys amb un 99'93%. Això és, sempre i quan el procés romangui sense alteracions i el pacient no pateixi canvis de tractaments, mori, etc.

Per una altra banda, si ha patit alguna, amb més d'un 96% (depenent del nombre d'arítmies tingudes) el pacient tindrà més de vint passat un any i amb un 100% passat els tres anys. Cal recordar que aquestes interpretacions es recolzen en les estimacions puntuals.

Per tal de calcular les estimacions per interval de les probabilitats, hem fet les potències a les matrius $P_{(s)}$ (matriu de probabilitats de transició de la mostra s obtinguda amb el mètode de Jackknife). Així doncs, tenim 59 estimacions de p_{ij}^{12} i de p_{ij}^{36} amb les que podem calcular la mitjana i la desviació tipus (amb la modificació que requereix el mètode de Jackknife) i fer els intervals de confiança com hem fet amb les probabilitats p_{ij} :

$$\begin{aligned}
 & \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 \\ (12) = & \begin{pmatrix} (-0'0304;0'1417) & (-0'0132;0'0707) & (-0'0168;0'0534) & (-0'0255;0'1190) & (-0'0094;0'1277) & (0'5906;0'9923) \\ & (-0'0102;0'0148) & (-0'0110;0'0157) & (-0'0380;0'0628) & (-0'0469;0'0958) & (0'8251;1'0919) \\ & & (-0'0005;0'0006) & (-0'0134;0'0158) & (-0'0286;0'0390) & (0'9457;1'0416) \\ & & & (-0'0037;0'0042) & (-0'0077;0'0109) & (0'9855;1'0108) \\ & & & & (-0'0005;0'0007) & (0'9993;1'0005) \\ & & & & & (1'0000;1'0000) \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix} \\
 & \\
 & \begin{matrix} & 0 & 1 & 2 & 3 & 4 & 5 \\ (36) = & \begin{pmatrix} (-0'0002;0'0005) & (0'0000;0'0002) & (-0'0002;0'0003) & (-0'0001;0'0005) & (0'0000;0'0005) & (0'9983;1'0002) \\ & (0'0000;0'0000) & (0'0000;0'0000) & (0'0000;0'0000) & (0'0000;0'0000) & (1'0000;1'0000) \\ & & (0'0000;0'0000) & (0'0000;0'0000) & (0'0000;0'0000) & (1'0000;1'0000) \\ & & & (0'0000;0'0000) & (0'0000;0'0000) & (1'0000;1'0000) \\ & & & & (0'0000;0'0000) & (1'0000;1'0000) \\ & & & & & (1'0000;1'0000) \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}
 \end{aligned}$$

Tanmateix, aquestes matrius quedaran de la següent manera si posem els límits a les probabilitats:

$$P^{(12)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} (0'0000;0'1417) & (0'0000;0'0707) & (0'0000;0'0534) & (0'0000;0'1190) & (0'0000;0'1277) & (0'5906;0'9923) \\ & (0'0000;0'0148) & (0'0000;0'0157) & (0'0000;0'0628) & (0'0000;0'0958) & (0'8251;1'0000) \\ & & (0'0000;0'0006) & (0'0000;0'0158) & (0'0000;0'0390) & (0'9457;1'0000) \\ & & & (0'0000;0'0042) & (0'0000;0'0109) & (0'9855;1'0000) \\ & & & & (0'0000;0'0007) & (0'9993;1'0000) \\ & & & & & (1'0000;1'0000) \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

$$P^{(36)} = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{pmatrix} (0'0000;0'0005) & (0'0000;0'0002) & (0'0000;0'0003) & (0'0000;0'0005) & (0'0000;0'05) & (0'9983;1'0000) \\ & (0'0000;0'0000) & (0'0000;0'0000) & (0'0000;0'0000) & (0'0000;0'0000) & (1'0000;1'0000) \\ & & (0'0000;0'0000) & (0'0000;0'0000) & (0'0000;0'0000) & (1'0000;1'0000) \\ & & & (0'0000;0'0000) & (0'0000;0'0000) & (1'0000;1'0000) \\ & & & & (0'0000;0'0000) & (1'0000;1'0000) \\ & & & & & (1'0000;1'0000) \end{pmatrix} & \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} \end{matrix}$$

3.3.6. Taula resum

A continuació es presenta una taula (*taula 4*) amb algunes de les estimacions de les probabilitats a un més, a un any i a tres anys. La seva lectura és directa ja que s'indica el nombre actual d'arítmies en un mes determinat i el nombre que es té passats els temps anteriors.

Probabilitats de transició a	Nº actual d'arítmies	Cap arítmia Addicional	Una arítmia addicional o més
1 mes	0	78'60% (68'70% ; 88'50%)	21'40% (11'50% ; 31'30%)
	1	60'00% (33'50% ; 86'50%)	40'00% (13'50% ; 66'50%)
	2	33'00% (0'00% ; 100'00%)	67'00% (0'00% ; 100'00%)
1 any	0	5'56% (0'00% ; 14'17%)	94'44% (85'83% ; 100'00%)
	1	0'22% (0'00% ; 1'48%)	99'78% (98'52% ; 100'00%)
	2	0'00% (0'00% ; 0'06%)	100'00% (99'94% ; 100'00%)
3 anys	0	0'02% (0'00% ; 0'05%)	99'98% (99'95% ; 100'00%)
	1	0'00% (0'00% ; 0'00%)	100'00% (100'00% ; 100'00%)
	2	0'00% (0'00% ; 0'00%)	100'00% (100'00% ; 100'00%)

Taula 4. Taula resum de les estimacions de les probabilitats

CAPÍTOL 4

CONCLUSIONS I DISCUSSIONS

4.1. CONCLUSIONS DE L'ESTUDI

4.1.1. Interpretació de les probabilitats de tenir arítmies

Després de realitzar l'estudi que desenvolupa el nostre projecte, podem concloure que existeix un increment del risc de tenir una futura arítmia a mesura que un pacient té més episodis.

Volíem veure si el fet de tenir una arítmia feia augmentar la probabilitat de tenir-ne una altre. Hem pogut observar que com més episodis tingui un pacient en un mes, és més possible tenir un altre en el següent. Aquesta suposició queda confirmada amb els resultats que obtenim a la matriu de probabilitats. Així trobem que la probabilitat de tenir alguna arítmia en el mes següent si no s'ha tingut cap és del 21'4%, si s'ha tingut una és del 40% i si s'han tingut dos és del 67%. Per tant, veiem com aquestes probabilitats augmenten a mesura que augmenten el nombre d'episodis que es tenen en un mes determinat.

Ara bé, si en un mes s'han tingut de tres a sis arítmies, la probabilitat de patir alguna altre és del 50%, i que per tant no augmenta. A partir d'aquí el procés és manté constant.

Tot i així, cal tenir present que tots els intervals de les estimacions de les probabilitats són molt amples, fins al punt de què puguin ser qualsevol valor ja que ens donen límits entre 0 i 1, la qual cosa vol dir que no tenim cap informació. Això és degut a que tenim una grandària mostral de 59, amb el que no es poden fer estimacions més exactes per algunes transicions.

Pel que fa a les probabilitats a llarg termini, la conclusió més important que en podem treure és que un pacient que no ha acumulat cap arítmia, després d'un any haurà acumulat més de vint amb una probabilitat del 79'26%, i després de tres anys amb un 99'93%. Això és, sempre i quan el procés romangui sense alteracions i el pacient no pateixi canvis de tractaments, mori, etc.

Per una altra banda, si ha patit alguna, amb més d'un 96% (depenent del nombre d'arítmies tingudes) el pacient tindrà més de vint passat un any i amb un 100% passat els tres anys. Cal recordar que aquestes interpretacions es recolzen en les estimacions puntuals.

4.1.2. Evolució del temps entre episodis

Un altre punt important en el nostre projecte ha estat estudiar el temps que passa des de que es produeix una arítmia fins que apareix la següent.

S'ha pogut observar que a partir de l'aparició del primer episodi, el temps entre arítmies disminueix fins que apareix la segona però que, a partir d'aquí es manté constant. Així doncs, podem dir que el procés és més ràpid després de la primera arítmia, i que a partir de la segona s'estabilitza.

4.2. DISCUSSIONS

Tots els resultats obtinguts són empírics i basats en els nostres dades. Hem de tenir present que comptem amb un nombre reduït de pacients i que, per tant les conclusions poden no ser tant precises com voldríem.

Cal fer notar que el nostre estudi és un anàlisi observacional i que té un enfocament purament predictiu. Així doncs, no s'estableixen relacions causals, com que "el fet de presentar una arítmia accelera el procés", tan sols es descriu la progressió de les arítmies.

Donat que l'aparell té un límit de capacitat de memòria, no es té constància de totes les arítmies que ha tingut un pacient a llarg del període de seguiment (si aquest ha tingut més d'un cert nombre d'arítmies). Per tant, pensem que les probabilitats de transició podrien estar infra-estimades.

Cal tenir present que els dos estudis tractats han seguit camins molt diferenciats. En el primer, hem estudiat el nombre d'arítmies per pacient, en el segon, hem tractat els temps entre aquestes. Tot i així, hem pogut veure que al final els resultats d'aquests anàlisis s'han complementat per a obtenir el següent: aquells pacients que tenen una arítmia en un mes, tenen la segona amb més probabilitat i més ràpidament. Un cop a passada la segona, la progressió de les arítmies resta constant, sense alteracions de temps. Tot i així, les conclusions de l'estudi del temps entre arítmies podrien limitar els resultats de les cadenes de Markov que combinen els canvis d'estat del primer mes al segon, amb els canvis del segon fins al tercer.

Una manera de contrastar aquest model seria estudiar la situació d'aquells pacients que tinguin un seguiment de 36 mesos. Malauradament, tant sols es disposa de 22 observacions que no poden considerar-se representatives del total de la mostra ja que podria ser que aquells individus seguits menys temps siguin els mes greus.

Respecte a l'article (*"An application of Markov models in estimating transition probabilities for postmenopausal women with osteoporosis"*, Zhengqing Li i Simon Pack. *Drug Information Journal*, Vol. 38, pp.41-46, 2004) que ha portat la idea original, volem remarcar les següents diferències. L'espai temporal tractat en la modelització de les fractures vertebrals són anys i nosaltres treballem amb mesos. Pel que fa a la creació de la cadena de Markov, a l'article es tenen en compte les configuracions (n_0, n_1, n_2, n_3) , en canvi, nosaltres no considerem el període n_0 , ja que no tenim l'històric dels pacients. D'altra banda, a l'article no es comptabilitzen les configuracions intermèdies, per exemple, la configuració $(0,1,1,2)$ només es té en compte en la transició $1 \rightarrow 2$ i no en la $1 \rightarrow 1$ com fem nosaltres, ja que observem una transició que es faci en dos mesos consecutius. Un altre punt és que a l'estudi de les fractures vertebrals es disposa d'un grup de control i un de tractats. Nosaltres no disposem de dades que ens permetin fer comparacions entre tractaments. Finalment, l'article utilitza Bootstrap per construir els intervals de confiança i nosaltres hem fet servir Jackknife.

El mètode d'estimació autosuficient (Bootstrap) és, com el Jackknife, un mètode per a obtenir la precisió d'un estimador de forma aproximada. Aquest mètode es basa en calcular directament la variança del estimador considerant la mostra com si fos la població. Així, es van prenent mostres a l'atzar amb reemplaçament amb les que s'estimen les probabilitats p_{ij} . Totes aquestes estimacions ens permetran calcular els intervals de confiança. Cal notar que, en el Jackknife les mostres diferien entre elles en un pacient, i en el Bootstrap podem tenir mostres que, per atzar, siguin iguals. En el moment en el que es necessita fer el càlcul de \hat{p}_{ij} només per les n mostres, el Jackknife comportarà càlculs més fàcil si n és més petita que 100. Malgrat tot, si observem les n mostres del Jackknife podem dir que usa una informació limitada sobre el estadístic \hat{p}_{ij} i per això podem tenir casos en el que el Jackknife pot ser menys eficient que el Bootstrap. Tot així el mètode del Jackknife es pot veure com a una aproximació al mètode de Bootstrap.

4.3. ESTUDIS FUTURS

Sobredispersió

Com ja hem observat al fer la descriptiva de la variable nombre d'episodis (*numepi*), hem trobat molta sobredispersió a les dades. Hem tractat d'anar més enllà estudiant la variable taxa d'episodis (*taxaepi*), però aquesta segueix estant sobredispersada.

Com que el nostre projecte no té com objectiu aquest tema, hem decidit deixar-ho de banda i aprofundir més en l'estudi del desenvolupament de les arítmies a llarg del temps. Tot i així, presentarem un petit esquema del què s'hauria de fer per tractar la sobredispersió.

Com hem comentat abans, tenim la hipòtesi de què la variable nombre d'episodis segueix una distribució de Poisson, ja que per definició aquesta variable compta el nombre d'arítmies que pot patir un pacient al llarg d'un període de seguiment. Per tal de comprovar-ho hauríem d'ajustar la variable per una Poisson de paràmetre λ , nombre mig de successos. Amb això, s'haurien de tenir en compte els següent punts:

1. Estimar el paràmetre λ a partir de les dades
2. Comparar la distribució empírica ($P(X < x_i)$) amb la teòrica Poisson (λ), és a dir, $P(X < x_i | \text{Poiss}(\lambda))$
3. Construir la recta de Henry (distribució teòrica a les ordenades i distribució empírica a les abscisses) o calcular la distància de Kolmogorov-Smirnov, per tal de comprovar si l'ajust és correcte.

Una altre hipòtesi, és ajustar la variable a una Binomial Negativa, que és el mateix que una mixtura de Poissons on els paràmetres es distribueixen segons una Gamma. En aquest cas, s'haurien d'estimar els paràmetres pel Mètode dels Moments.

Un cop trobada la distribució que s'ajusta millor les nostres dades, caldria fer una simulació per tal de verificar els resultats.

Simulacions

A continuació presentarem una altra possibilitat per a estudiar les dades que hem fet servir en el nostre projecte.

L'objectiu en aquest apartat era aconseguir simular el nombre d'episodis que han tingut els pacients al llarg del temps que han estat en seguiment, per tal de poder explicar el que caracteritza la progressió del temps del ritme cardíac dels pacients.

Per portar a terme aquest estudi, en primer lloc, hauríem hagut d'estudiar la distribució que segueix la variable nombre d'episodis, i poder determinar els seus paràmetres.

Hem començat plantejant dues simulacions, una continua i una altra discreta, que generaran per a cada pacient un recompte del nombre d'episodis en un temps de seguiment determinat. La primera, *simulacont*, genera mitjançant la distribució, determinada a priori per les dades de la mostra, els temps entre episodis i comparar els temps acumulats en el període de seguiment. La segona, *simuladisc*, genera mitjançant la Distribució Uniforme valors que es comparen amb la probabilitat de tenir o no arítmies.

Simulació Continua (Annex 5)

Simulacont genera per a cada pacient i ($i: 1 \dots NP=59$), un comptador *compt* (inicialitzat a zero) i un valor (x) de la distribució determinada a partir d'un estudi previ. En aquest cas, podem triar com exemple la Distribució Exponencial, on el seu paràmetre està estimat a partir de les dades (%L) i és la taxa d'episodis:

$$\bar{x}/TT = 80'6441/56 = 1'44 \approx 1'5 \quad (Taula 11)$$

Aquest valor simula el temps que ha passat fins al primer episodi del pacient i .

Mentre que el valor generat x , sigui menor que el temps de seguiment TT , en aquest cas 56 mesos (*Taula 12*), el comptador s'incrementa en una unitat i es genera un altre valor (y) amb la distribució indicada, essent aquest el següent episodi. Tot seguit s'actualitza x com la suma dels dos valors generats anteriorment ($x = x + y$),

és a dir, el temps acumulat fins a la següent arítmia. Així, es compara una altra vegada el valor acumulat (x) amb el temps de tractament. Si la suma és més petita que el temps en qüestió (TT), llavors s'incrementa un altre cop el comptador. Si no, sortim del bucle, ja que ens indica que el temps acumulat fins a la següent arítmia és superior al temps de seguiment, és a dir, que no es poden tenir en compte més episodis.

Obtenim, d'aquesta manera, el nombre d'episodis que ha tingut el pacient i en el període de seguiment, indicat pel valor final del comptador. Iterant tantes vegades com quantitat de pacients, queda determinada finalment la variable nombre d'episodis.

A continuació presentem l'anàlisi necessari per tal de portar a terme la simulació *simulacont*:

- *Descriptiva de la variable nombre d'episodis (Taula 11)*

	N	Mínimo	Máximo	Media	Desv. típ.
NUMEPI	59	1,00	497,00	80,6441	116,79658
N válido	59				

- *Descriptiva del temps de seguiment en mesos (Taula 12)*

	N	Mínimo	Máximo	Media	Desv. típ.
Meses de seguimiento	57	0,27	56,00	28,7327	16,40964
N válido	57				

Simulació Discreta (Annex 6)

Simuladisc genera, per a cada pacient i ($i: 1....NP=59$) valors de la Distribució Uniforme $[0,1]$ i els compara amb la probabilitat de què un malalt pugui tenir una arítmia ($\%PI = 0,5$). Es crea també un comptador *compt* (inicialitzat a zero), que emmagatzemarà el nombre d'episodis per a cada pacient i .

Si aquests valors generats són més petits que la probabilitat indicada, llavors el comptador augmenta en una unitat. Si no, surt del bucle i comença a iterar pel següent pacient, i així fins a completar els 59 valors de la variable nombre d'episodis.

L'objectiu d'aquestes simulacions és poder comparar si les nostres dades s'ajusten a les generades de la variable nombre d'episodis. Hauria estat interessant de

continuar aquest estudi, afegint per exemple, els tipus d'arítmies (comentades en capítols anteriors) i veure com varia el nombre d'episodis, i afegint el tipus de tractaments i veure també si disminueixen o no el nombre d'arítmies.

Cal comentar que al fer la descriptiva de la variable nombre d'episodis, ens trobàvem amb molta sobredispersió en les dades, ja que la mitjana és molt més petita que la desviació tipus: $116,797^2 \gg 80,644$. Aquest fet, ha estat comentat en el punt 3.1.

Altres idees futures

Es podrien realitzar altres estudis incorporant factors importants al model, el càlcul i la interpretació seran més complicats. En cas que l'objectiu fos verificar l'eficàcia de les teràpies s'hauria de estudiar les variables disponibles a la mostra, ja que faria falta un bon disseny per tal de tenir les variables confusoras controlades i per aprofitar tota la informació de què disposem al DAI. Aquest anàlisi necessita una investigació més exhaustiva, ja que s'haurien de considerar diferents tipus d'arítmies, diferents tipus de teràpies i també la duració dels episodis.

Pel que fa als tipus d'arítmies auriculars, taquicàrdia auricular (TA) i fibril·lació auricular (FA), es poden tenir interessos diferents si s'estudien juntes o separades. L'anàlisi en conjunt es faria quan es vulgues veure la relació entre les arítmies auriculars i les ventriculars i separat si s'esperessin resultats diferents que no es poguessin barrejar.

Altre anàlisi interessant podria ser estudiar l'eficiència de detecció del DAI, ja que a vegades, la detecció no es correspon amb el electrograma que fa el metge i subministra teràpies més fortes del què el malalt necessita en aquest moment.

Una altra aspecte a tenir en compte, és que els pacients han estat seguits temps diferents. Al fer l'estudi del nombre d'episodis per pacient podríem utilitzar tècniques d'anàlisi de supervivència per tal de poder considerar tots els pacients amb el seus temps de seguiment.

BIBLIOGRAFIA

- (1) Daniel Peña Sánchez de Rivera. *Estadística: modelos y métodos (1. Fundamentos)*. Alianza editorial. Madrid 1986
- (2) Bradley Efron, Robert J. Tibshirani. *An introduction to the Bootstrap*. Chapman & Hall. United States of America 1993
- (3) Mark j. Schervish. *Theory of Statistics*. Springer. United States of America 1997
- (4) Jun Shao, Dongsheng Tu. *The Jackknife and Bootstrap*. Springer. United States of America 1996
- (5) Bradley Efron. *The Jackknife, the Bootstrap and other Resampling Plans*. Society for Industrial and Applied Mathematics. Philadelphia, Pennsylvania 1982
- (6) Bo-Cheng Wei. *Exponential Family. Nonlinear Models*. Ed. Springer. Singapore 1998
- (7) Alan Stuart, Keith Ord. *Kendall's advanced theory of statistics. Volume 1: Distribution theory 6th edition*. Ed. Edward Arnold. London 1994
- (8) Norman L. Johnson, Samuel Kotz, Adrienne W. Kemp. *Univariate Discrete Distributions. Second edition*. Ed. John Willey and Sons. New York 1993
- (9) Esteve Codina i Lúdia Montero. *Apunts d'Investigació Operativa Estocàstica (Cadenes de Markov)*. Barcelona 2000
- (10) Redacció Konrad Diem i Cornelius Lentner. *Tablas científicas. Séptima edición*. Documenta Geigy. Barcelona 1975.

(11) D. Collet. *Modeling survival data in medical research*. Ed. Chapman & Hall. London 1994

(12) David G. Kleinbaum. *Survival Analysis. A self-learning text*. Ed. Springer. New York 1995

(13) Ryan-Joiner. *Minitab handbook*. Ed. duxbury press. California 1994

(14) http://www.texasheartinstitute.org/arrhy_sp.html

(15) E. Cobo, S. Videla, B. Clotet, M.L. Calle, C. Corchero, G. Gómez, V. De Gruttola, K. Langohr. *Investigación clínica y estadística: una visión multidisciplinar con aplicaciones en estudios de VIH/Sida*. Ed. Fundació de Lluita contra la SIDA. Badalona 2003.

(16) A. Prat, X.Tort-Martorell, P. Grima, L. Pozueta. *Métodos estadísticos*. Ed. UPC. Barcelona 1997.

ANNEX 1: Article

STATISTICS 41

An Application of Markov Models in Estimating Transition Probabilities for Postmenopausal Women with Osteoporosis

Zhengqing Li, PhD

Section Head, Biometrics
and Statistical Sciences,
Procter & Gamble
Pharmaceuticals, Mason,
Ohio

Simon Pack, PhD

Associate Director, New
Drug Development, Europe,
Procter & Gamble
Pharmaceuticals,
United Kingdom

Vertebral fracture is a common consequence of osteoporosis for postmenopausal women. Modeling the progression of the disease in vertebral fractures has been of great interest to physicians as well as patients. In this paper, we propose a simple Markov chain model to study the progression of the vertebral fractures for untreated postmenopausal women with osteoporosis. In this model, the state space consists of $\{0, 1, 2, \dots, 13\}$, which represents the possible number of vertebral fractures for a patient, as commonly assessed via spinal x-ray. Based on the model, a

simple estimator of the one-year transition probability matrix is proposed. An estimate for the m -year transition probability matrix is then derived following an application of the Chapman-Kolmogorov equations. To estimate the confidence intervals for the m -year transition probabilities, a bootstrap procedure is proposed. The proposed bootstrap procedure can also be employed to construct confidence intervals for differences between two treatment groups. The methods are illustrated with data from clinical trials for an anti-resorptive therapy.

Key Words

Bootstrap; Markov chains;
Osteoporosis; Stochastic
processes; Transition
probabilities

Correspondence Address

Zhengqing Li, Biometrics
and Statistical Sciences, The
Procter & Gamble Company,
5700 Mason-Montgomery
Rd., Box 2199, Mason, OH
45040; e-mail: li_zh@pg.com

INTRODUCTION

Vertebral fractures, that is, severe deformation of spinal vertebra, are a well-recognized consequence of postmenopausal bone loss and are the most common osteoporotic fracture (1). All vertebral fractures are associated with increased mortality and morbidity, including back pain and decreased activity. Previous data have indicated that prevalent vertebral fractures substantially increase the risk of future vertebral fractures (2). However, the speed of progression of the disease has not been well characterized and is usually addressed via epidemiological studies. Data from clinical trials of osteoporosis therapies provide unique data that facilitate alternative modeling techniques.

As a useful mathematical tool, Markov models have been well developed in the mathematical and statistical literature (3) and have been used in many areas (4). For example, for aggregated data (data that consist of the number of individuals in each state at specified observation times), statistical methods have been developed in estimating the transition probabilities (5).

In the therapeutic area of postmenopausal osteoporosis, vertebral fractures have been assessed in clinical trials of up to three years dura-

tion using annual spinal x-rays to see whether a patient sustains a fracture and the time period when the fracture occurs (6, 7, 8). It is important to note that patients may experience multiple vertebral fractures in any time period. Typically, assessment is made on 15 vertebrae in the thoracic and lumbar regions of the spine.

In this paper, we demonstrate how to model the progression of vertebral fractures using a simple Markov model for this type of data. Based on this Markov model, we estimate the one-year transition probability matrix using a simple estimate. The long-term transition probability matrix is derived using matrix multiplication with the Chapman-Kolmogorov equations. To construct confidence intervals for m -year long-term transition probabilities, we propose using a bootstrap method that preserves the correlation structure among the transition probabilities for a Markov chain. The proposed bootstrap method is also adopted to construct confidence intervals for between-group differences.

STATISTICAL MODEL

We consider a stochastic process $\{X_n\}$ that takes on a finite or countable number of possible values. We denote the possible values of the

process by a set of nonnegative integers $[0, 1, 2, \dots, 13]$. In the example that we described for osteoporosis, the nonnegative integers represent the number of vertebral fractures that a patient may sustain. For example, if $X_n = j$, then a patient is said to have j vertebral fractures at time n . We suppose that whenever the process is in state i , there is a fixed probability p_{ij} that it will next be in state j ; that is, we suppose that for all $n \geq 0$:

$$P\{X_{n+1} = j | X_n = i\} = p_{ij}.$$

Since a fracture stays forever once it occurs, we have $p_{ji} = 0$ for $j < i$. In this model, we assume the conditional distribution of any future state X_{n+1} depends only on the present state and is independent of the past states. In the example for osteoporosis, this is equivalent to the assumption that the past history impacts the risk of future vertebral fractures only through the current number of vertebral fractures. Let $\mathbf{P} = [p_{ij}]$ denote the one-step transition probability matrix. We now define the m -step transition probability $p_{ij}^{(m)}$ to be the probability that a process in state i will be in state j after m additional steps, that is:

$$P\{X_{n+m} = j | X_n = i\} = p_{ij}^{(m)}.$$

Following the Chapman-Kolmogorov equations (9), the m -step transition probability matrix, $\mathbf{P}^{(m)} = [p_{ij}^{(m)}]$, can be obtained by multiplying the matrix \mathbf{P} by itself m times, that is:

$$\mathbf{P}^{(m)} = \mathbf{P}^m.$$

In the example for osteoporosis, the m -step transition probability, $p_{ij}^{(m)}$, represents the probability that a patient will sustain j vertebral fractures in m years given that she has i vertebral fractures now.

ESTIMATES OF THE TRANSITION PROBABILITIES

We consider the following data collected from a clinical study with a duration of three years. Suppose vertebral fracture assessment is made for each patient at baseline, end of year 1, end of year 2, and end of year 3. Denote the cumulative number of vertebral fractures at the four

time points by n_0, n_1, n_2 , and n_3 . The fracture status path for a patient during the study can be represented by the configuration (n_0, n_1, n_2, n_3) . Clearly, the numbers of vertebral fractures satisfies $0 \leq n_0 \leq n_1 \leq n_2 \leq n_3 \leq 13$ since a fracture will not disappear and since a patient only has 13 vertebrae.

To calculate the one-year transition probabilities, we need to determine the paths for each transition. A requirement for the transition probabilities from state i is $\sum_j p_{ij} = 1$. To fulfill this requirement, the configurations for transitions from i to k and i to l must be mutually exclusive for $k \neq l$. For all transitions starting from zero fracture, the configuration space can be represented by $\Omega_0 = \{(0, n_1, n_2, n_3); 0 \leq n_1 \leq n_2 \leq n_3 \leq 13\}$. For a specific transition from zero to j , we have the following configurations.

Transitions	Configurations
0 to 0	(0, 0, 0, 0)
0 to 1	(0, 1, *, *), (0, 0, 1, *), (0, 0, 0, 1)
0 to 2	(0, 2, *, *), (0, 0, 2, *), (0, 0, 0, 2)
0 to 3	(0, 3, *, *), (0, 0, 3, *), (0, 0, 0, 3)
0 to j	(0, j , *, *), (0, 0, j , *), (0, 0, 0, j)

In the configurations above, we use * to denote any possible nonnegative integer that is greater than or equal to the integer in the previous year. Since the transition from zero to zero corresponds to a path of no change in fracture status, the only configuration for this transition is (0, 0, 0, 0). The configurations such as (0, 0, 0, 1) and (0, 0, 1, 1) correspond to paths in which the fracture status has changed. Therefore, these configurations do not belong to the transition from zero to zero. Including these configurations in the transition from zero to zero also violates the mutual exclusivity in the transition probability calculation. For the transitions starting from one fracture, the configuration space can be written as $\Omega_1 = \{(1, *, *, *), (*, 1, *, *), (*, *, 1, *)\}$. Since the transition from one to one corresponds to paths in which a patient keeps her fracture status at one after she sustains the first fracture, the configurations associated with these paths can be written as $\{(*, *, 1, 1)\} = \{(0, 0, 1, 1), (0, 1, 1, 1), (1, 1, 1, 1)\}$. A configuration such as (0, 1, 1, 2) does not belong to the transi-

from one to one since the fracture status is changed after the first fracture occurs. Specifically, we can write the configurations explicitly for the transition from one to j .

Transitions	Configurations
1 to 0	Not applicable
1 to 1	(0, 0, 1, 1), (0, 1, 1, 1), (1, 1, 1, 1)
1 to 2	(1, 2, *, *), (*, 1, 2, *), (*, *, 1, 2)
1 to 3	(1, 3, *, *), (*, 1, 3, *), (*, *, 1, 3)
1 to j	(1, j , *, *), (*, 1, j , *), (*, *, 1, j)

Following the same method, we can list the configurations for the transition from i fractures to j fractures. Note it is impossible to make the transition from i to j for $j < i$ since once a fracture occurs, it will stay forever.

Let k_{ij} denote the number of patients whose paths belong to the transition from i to j . The one-step transition probability from i to j can be estimated by:

$$\hat{p}_{ij} = \frac{k_{ij}}{\sum_k k_{ik}} \quad \text{for } i, j = 0, \dots, 15, j \geq i.$$

In the calculation above, patients are used only once for the transitions with the same starting state due to the mutual exclusivity of counting configurations. However, patients could be counted multiple times when calculating the transition probabilities with different starting states. For example, the configuration (0, 0, 1, 1) is counted for both the transition from zero to one and the transition from one to one. Therefore, the estimates of the transition probabilities with different starting states are not independent. Based on the one-year transition matrix estimated above, the m -year transition probability matrix can be calculated by applying the Chapman-Kolmogorov equations through matrix multiplication. Let $\hat{p}_{ij}^{(m)}$ denote the estimate of the m -year transition probability from i to j . The m -year transition probability from zero fracture to one or more fractures was calculated as:

$$\hat{p}_{0j}^{(m)} = \sum_{i=0}^{15} \hat{p}_{0i} \hat{p}_{ij}^{(m-1)}.$$

Similarly the m -year transition probability from zero to two or more fractures can be calculated as:

$$\hat{p}_{0j}^{(m)} = \sum_{i=0}^{15} \hat{p}_{0i} \hat{p}_{ij}^{(m-1)}.$$

To calculate the 95% confidence intervals for m -year transition probabilities, we applied a bootstrap method for Markov chains (10). This method was used because we want to preserve the correlation structure of the matrix. Specifically, the following algorithm can be implemented:

1. Let N be the total number of patients used in the calculation. Select patients randomly, with replacement, N times from the patient population and denote this sample by S_1 .
2. Based on the sample S_1 , we calculate the m -year transition probabilities following the procedures described.
3. Replicate (1) and (2) M times and generate the bootstrap sampling distribution for each transition probability.
4. The mean and standard error can be calculated based on the bootstrap samples generated in (3). Based on the bootstrap samples, one can calculate the 95% CI using the 2.5% and 97.5% percentiles.

Clearly, the above algorithm preserves the correlation structure of the transition matrix.

While the primary interest of this research was to model the progression of the disease for placebo-treated patients, it was also desirable to make between-treatment comparisons. Under the framework of bootstrap sampling outlined, there is no essential difficulty with such between-group comparisons. Let p_i and p_i^* be the bootstrap estimates (means of the bootstrap samples) of a quantity of interest from the m -year transition matrices for the placebo and treatment groups, respectively. Let SE_i and SE_i^* be the standard errors calculated from the bootstrap samples for the two estimates. Since the patients from the two treatment groups are independent, the 95% two-sided CI for the difference, $p_i - p_i^*$, can be calculated as:

$$p_i - p_i^* \pm 1.96 \sqrt{SE_i^2 + SE_i^{*2}}.$$

AN EXAMPLE

Risedronate is a pyridinyl bisphosphonate with high affinity for hydroxyapatite crystals in bone. The efficacy of risedronate in reducing the risk

of vertebral fractures compared to placebo in postmenopausal women with osteoporosis has been demonstrated in two randomized, double-blind, placebo-controlled studies (6,7). In these two studies (VERT NA and VERT MN), the primary efficacy endpoint was the incidence of vertebral fractures over a three-year treatment period. In addition, risedronate has also been shown to reduce the risk of hip fractures in elderly women in a large randomized and placebo-controlled study (8). In all these studies, vertebral fracture assessments were made at baseline and also annually over the three-year treatment period.

In this analysis, we estimated the long-term transition probabilities in vertebral fractures for the placebo patients who had osteoporosis from the three studies (VERT NA, VERT MN, and HIP) using the methods described in previous sections. The placebo patients in all studies received calcium or vitamin D supplementation. The population that we were interested in was those osteoporotic patients as assessed by femoral neck t scores ≤ -2.5 or prevalent vertebral fractures. A total of 2526 placebo patients were included in the analysis. A 14×14 one-step transition matrix was generated based on the data from the 2526 patients. In this matrix, all transition probabilities p_{ij} equal to zero for $j < i$. It also appears that patients were less likely to make a transition to four or more new fractures within a year. As a consequence, the number of configurations, k_{ij} , was zero for most cells in which $j > i + 5$. Based on the 1-year transition

matrix we derived the 5-year and 10-year transition matrices. For illustration purposes, we were interested in the following transition probabilities:

- The probabilities that a patient will sustain 1 or more or 2 or more vertebral fractures in the next 5 or 10 years given that she has no prevalent vertebral fracture.
- The probabilities that a patient will sustain 1 or more or 2 or more vertebral fractures in the next 5 or 10 years given that she has 1 prevalent vertebral fracture.
- The probabilities that a patient will sustain 1 or more or 2 or more vertebral fractures in the next 5 or 10 years given that she has 2 prevalent vertebral fractures.

These probabilities provide an assessment of how the disease progressed when patients were not treated with active therapies. In the analysis, all placebo patients with postmenopausal osteoporosis who had known fracture status at baseline and had postbaseline vertebral fracture assessments were included. When calculating the 95% confidence intervals, 500 bootstrap samples were generated. The estimated probabilities and the corresponding 95% confidence intervals are included in Table 1. For an untreated woman who had no vertebral fractures, the probability that she will sustain at least 1 vertebral fracture was estimated to be approximately 33% in 5 years and 55% in 10 years. For those patients who have sustained fractures, the risk that they will sustain more fractures in 5 or 10 years increased dramatically. This analysis clear-

TABLE 1

Estimated Long-Term Transition Probabilities (in 5 and 10 years) and the Corresponding 95% Confidence Intervals

Number of Present Fractures	Transition Probability in 5 Years		Transition Probability in 10 Years	
	To Additional 1 or More Fractures	To Additional 2 or More Fractures	To Additional 1 or More Fractures	To Additional 2 or More Fractures
0 (N = 703)	33.1% (25.5%, 40.6%)	11.4% (7.7%, 15.8%)	55.1% (44.5%, 64.7%)	29.0% (21.9%, 37.0%)
1 (N = 567)	46.7% (37.9%, 55.5%)	17.2% (12.5%, 22.5%)	71.4% (61.5%, 80.2%)	42.4% (34.1%, 51.2%)
2 (N = 407)	54.8% (44.9%, 64.4%)	26.5% (20.4%, 33.3%)	79.3% (69.7%, 87.3%)	57.0% (47.9%, 66.3%)

ly indicates the importance of treating patients early and the necessity of preventing the first vertebral fracture.

To illustrate how one can compare the transition probabilities between the two treatment groups, we also calculated the transition matrix for the risedronate 5 mg group. For a risedronate 5 mg treated woman with no prevalent vertebral fractures, the probability that she will sustain at least one vertebral fracture in five years was estimated to be 19.9%. This probability was 15.2% (95% CI: 5.3%, 23.1%), lower than a similar placebo-treated patient and associated with a relative risk of 0.60. The probability of sustaining at least two new fractures in five years for a risedronate-treated patient was estimated to be 5.5%, 5.9% (95% CI: 1.1%, 10.7%) lower than a similar placebo-treated patient. The relative risk associated with this difference was 0.48. These differences were statistically significant since zero was not included in the 95% CIs for the differences in transition probabilities.

CONCLUSION AND DISCUSSION

Although the application of Markov models has been investigated in many different areas (4), to our knowledge they have not been used in trials of therapies for osteoporosis. When fracture assessments are made periodically, a Markov model provides a simple approach to modeling the progression of disease. The estimates of the long-term transition probabilities in fractures are not only instrumental to researchers for understanding the diseases but they also provide guidance to physicians when treating patients and to health policymakers and regulatory practitioners in public health policymaking. We proposed using the bootstrap method to calculate the corresponding confidence intervals for the long-term transition probabilities. With the current advancements in computing, this method is computationally straightforward and can be easily implemented. Furthermore, this method preserves the correlation structure of the transition matrix.

The primary interest of this research was to

provide a statistical approach to modeling the progression of osteoporosis for an untreated population. Under the framework of bootstrap sampling, one can also make between-group comparisons for transition probabilities. However, the between-group comparisons outlined should be viewed as hypothesis generating rather than as hypothesis testing. An appropriately designed randomized controlled study is the definitive tool for generating efficacy data for a new agent. The between-group differences that we presented in the example based on Markov models with bootstrap sampling should be interpreted with this caution. The risk reductions for patients without prevalent vertebral fractures based on our modeling, however, appear to be consistent with the risk reductions observed in randomized clinical trials (6,7). This provides further support to the validity of our methods in modeling the progression of the disease.

One assumption that we have made in the model is that all past history and other factors (eg, bone mineral density) of a patient affect the risk of future vertebral fractures only through the present number of vertebral fractures that the patient sustains, that is, given the present number of vertebral fractures that a patient sustains, the risk of future vertebral fractures is independent of all the past history. While the other factors such as bone mineral density may have an impact on the risk of fractures, analyses have shown that the number of prevalent vertebral fractures is the most important factor in determining the risk of future fractures (2). Although similar statistical approaches could be developed to incorporate other important factors in the model, the calculation and interpretation of the model will be more complicated. This is an analysis that deserves further investigation.

Acknowledgments — The authors would like to express their gratitude to Michael Bramley for his help with the SAS programming needed for the example. The authors are also thankful to Drs. Michael Meredith, Eileen King, and Mike Manhart for their valuable comments and suggestions.

REFERENCES

1. Ross PD. Clinical consequence of vertebral fractures. *Am J Med.* 1997;105(suppl):50s-47s.
2. Lindsay R, Silverman SL, Cooper C, Hanley DA, Barton I, Broy SB, Licata A, Benhamou I, Geusens P, Flowers K, Stracke H, Seeman E. Risk of new vertebral fracture in the year following a fracture. *JAMA.* 2001;285:520-525.
3. Cox DR, Miller HD. *The Theory of Stochastic Processes.* London, United Kingdom: Methuen; 1965.
4. Chiang CL. *An Introduction to Stochastic Processes and their Applications.* New York, NY: Krieger; 1980.
5. Kalbfleisch JD, Lawless JF, Vollermer WM. Estimation in Markov models from aggregate data. *Biometrics.* 1985;39:907-919.
6. Reginster JY, Minne HW, Sorensen OH, Hooper M, Rous C, Brandi ML, Land B, Ethgen D, Puck S, Roumagnac I, Eastell R. Randomized trial of the effects of risedronate on vertebral fractures in women with established postmenopausal osteoporosis. *Osteoporosis Int.* 2000; 11:87-91.
7. Harris ST, Watts NB, Genant HK, McKeever CD, Hangartner T, Keller M, Chesnut C, Brown J, Eriksen EL, Hossain MS, Axelrod DW, Miller PD. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis. *JAMA.* 1999;282:1344-1352.
8. McClung MR, Geusens P, Miller PD, et al. Effect of risedronate on the risk of hip fracture in elderly women. *New Engl J Med.* 2001;344:335-340.
9. Ross SM. *Introduction to Probability Models.* San Diego, CA: Academic Press; 1995.
10. Athreya KB, Fuh CD. Bootstrap Markov chains: countable case. *J Stat Plan Inference.* 1992; 53:311-351.

ANNEX 2: Càlculs de probabilitats

A continuació es presenta una taula amb les transicions possibles entre estats, les seves configuracions i el nombre de pacients per a cadascuna d'elles. A partir d'aquests, es calculen les estimacions de les probabilitats p_{ij} .

TRANSICIONS i j	CONFIGURACIONS (N1,N2,N3)	Kij (N1,N2,N3)	Kij	Estima Pij
0 0	(0,0,0)x2 (0,0,1) (0,0,2) (0,0,3) (0,0,4) (0,0,5)	24x2=48 3 1 0 2 1	55	55/70=0.786
0 1	(0,1,1) (0,1,2) (0,1,3) (0,1,4) (0,1,5) (0,0,1)	3 1 0 0 0 3	7	7/70=0.1
0 2	(0,2,2) (0,2,3) (0,2,4) (0,2,5) (0,0,2)	0 0 0 0 1	1	1/70=0.014
0 3	(0,3,3) (0,3,4) (0,3,5) (0,0,3)	2 1 0 0	3	3/70=0.043
0 4	(0,4,4) (0,4,5) (0,0,4)	0 1 2	3	3/70=0.043
0 5	(0,5,5) (0,0,5)	0 1	1	1/70=0.014
1 0	-	-	-	-
1 1	(1,1,1)x2 (1,1,2) (1,1,3) (1,1,4) (1,1,5) (0,1,1)	2x2=4 1 0 0 1 3	9	9/15=0.6
1 2	(1,2,2) (1,2,3) (1,2,4) (1,2,5) (0,1,2) (1,1,2)	1 1 0 0 1 1	4	4/15=0.266
1 3	(1,3,3) (1,3,4) (1,3,5) (0,1,3)	0 0 0 0	0	0

		(1,1,3)	0		
1	4	(1,4,4)	0	1	1/15=0.067
		(1,4,5)	1		
		(0,1,4)	0		
		(1,1,4)	0		
1	5	(1,5,5)	0	1	1/15=0.067
		(0,1,5)	0		
		(1,1,5)	1		
2	0	-	-	-	-
2	1	-	-	-	-
2	2	(2,2,2)x2	0	1	1/3=0.33
		(2,2,3)	0		
		(2,2,4)	0		
		(2,2,5)	0		
		(0,2,2)	0		
		(1,2,2)	1		
2	3	(2,3,3)	0	2	2/3=0.67
		(2,3,4)	0		
		(2,3,5)	1		
		(0,2,3)	0		
		(1,2,3)	1		
		(2,2,3)	0		
2	4	(2,4,4)	0	0	0
		(2,4,5)	0		
		(0,2,4)	0		
		(1,2,4)	0		
		(2,2,4)	0		
2	5	(2,5,5)	0	0	0
		(0,2,5)	0		
		(1,2,5)	0		
		(2,2,5)	0		
3	0	-	-	-	-
3	1	-	-	-	-
3	2	-	-	-	-
3	3	(3,3,3)x2	1x2=2	4	4/8=0.5
		(3,3,4)	0		
		(3,3,5)	0		
		(0,3,3)	2		
		(1,3,3)	0		
		(2,3,3)	0		
3	4	(3,4,4)	1	3	3/8=0.375
		(3,4,5)	1		
		(0,3,4)	1		
		(1,3,4)	0		
		(2,3,4)	0		
		(3,3,4)	0		
3	5	(3,5,5)	0	1	1/8=0.125
		(0,3,5)	0		
		(1,3,5)	0		
		(2,3,5)	1		
		(3,3,5)	0		
4	0	-	-	-	-
4	1	-	-	-	-

4	2	-	-	-	-
4	3	-	-	-	-
4	4	(4,4,4)x2 (4,4,5) (0,4,4) (1,4,4) (2,4,4) (3,4,4)	2x2=4 1 0 0 0 1	6	6/13=0.46
4	5	(4,5,5) (0,4,5) (1,4,5) (2,4,5) (3,4,5) (4,4,5)	3 1 1 0 1 1	7	7/13=0.54
5	0	-	-	-	-
5	1	-	-	-	-
5	2	-	-	-	-
5	3	-	-	-	-
5	4	-	-	-	-
5	5	(5,5,5)x2 (0,5,5) (1,5,5) (2,5,5) (3,5,5) (4,5,5)	3x2=6 0 0 0 0 3	9	9/9=1

ANNEX 3: Macro Minitab Jackknife

A continuació es presenta una macro per a Minitab amb la que ens hem ajudat per tal de fer les mostres pel mètode de Jackknife.

```

MACRO
MCONFIG N1 N2 N3 N11 N22 N33 K.1-K.NP /*paràmetres que se li passen*/

MCOLUMN N1 N2 N3 N11 N22 N33 K.1-K.NP /*declaració dels paràmetres*/
MCONSTANT I J CONT NP NC PE

LET NP = 59 /*nombre de pacients*/
LET NC = 106 /*nombre de configuracions*/

DO PE = 1:NP /*es fa les mostres on hi ha un pacient exclòs cada cop*/
DO I=1:NC /*es fan comparacions amb totes les configuracions teòriques*/
    LET CONT = 0
    DO J = 1:NP
        IF (J ~= PE) /*no es té en compte el pacient exclòs*/
            IF (N1(I) = N11(J) & N2(I) = N22(J) & N3(I) = N33(J))
                LET CONT = CONT +1
            ENDIF
            /*si la configuració observada és com la teòrica incrementem el
            comptador d'aquesta configuració*/
        ENDIF
    ENDDO
    LET K.PE(I) = CONT
ENDDO
ENDDO

ENDMACRO

```

ANNEX 4: Estudi complementari de l'evolució del temps entre arítmies

A continuació presentarem un estudi en el que s'han tingut present 10 pacients dels que no hi havia suficient informació per a tenir-les en compte en l'estudi de l'evolució del temps entre arítmies, de manera que se'ls hi han assignat el següents valors:

- temps que passa entre no tenir cap arítmia i tenir una (t_{01}) serà el temps de seguiment, ja que aquest és el temps mínim en el que es produirà una arítmia.
- suposarem el temps que passa entre tenir una arítmia i tenir dos (t_{12}) i el temps que passa entre tenir dos i tenir tres (t_{23}) arítmies és zero.

Pel que fa als 2 pacients restants no tenim constància del seu temps de seguiment, per tant no podem assignar-hi cap valor.

Temps entre arítmies

Observem els gràfics on es relacionen els temps de seguiments entre arítmies:

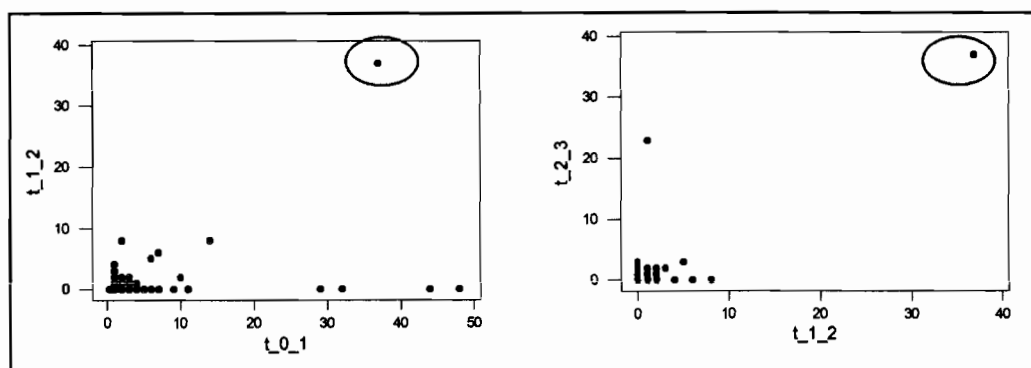


Fig.3. Plot dels temps entre arítmies.

La observació 20 distorsiona els gràfics, l'hem tret per la construcció dels gràfics, però l'hem inclòs en els càlculs numèrics.

Gràfics on es relacionen els temps entre arítmies, sense l'observació 20

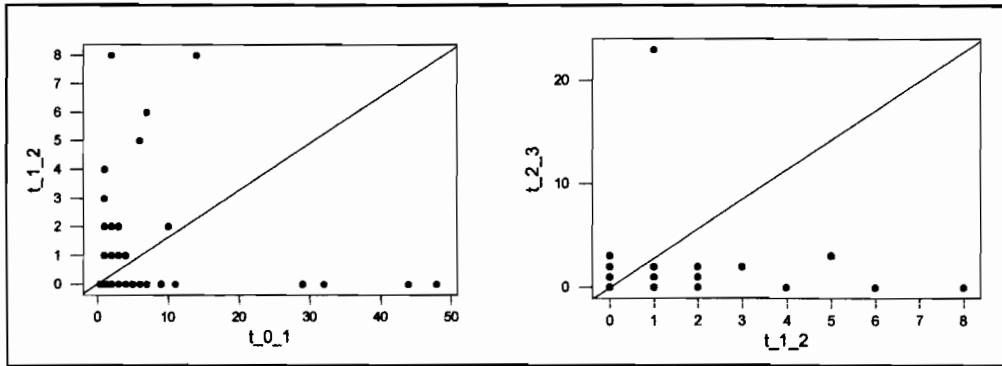


Fig.22. Plot dels temps entre arítmies.

Observem que a mesura que el temps entre les arítmies és major, la dispersió augmenta. La majoria d'observacions es troben en espais de temps curts entre arítmies.

Presentem la descriptiva de les variables:

Variable	N	N*	Mean	Median	TrMean	StDev
t_0_1	57	2	6,08	2,00	4,23	10,59
t_1_2	57	2	1,544	0,000	0,686	5,131
t_2_3	57	2	1,456	0,000	0,392	5,704

Variable	SE Mean	Minimum	Maximum	Q1	Q3
t_0_1	1,40	0,27	48,00	1,00	5,50
t_1_2	0,680	0,000	37,000	0,000	1,000
t_2_3	0,756	0,000	37,000	0,000	1,000

Fig.23. Descriptiva dels temps entre arítmies.

Observem resultats anàlegs als del primer estudi, veiem un decreixement molt important en la mitjana de t_{01} a la de t_{12} , prenent valors de 6'080 i 1'544. En canvi de t_{12} a t_{23} no es veu tal disminució, ja que les mitjanes són pràcticament iguals, prenent valors de 1'544 i 1'456 respectivament.

Per a comprovar que els temps es redueixen, hem realitzat la comparació de mitjanes d'aquestes variables mitjançant l'anàlisi de la varianza per a dades aparellades per pacient:

$$\begin{cases} H_0: \mu_{t_{01}} = \mu_{t_{12}} = \mu_{t_{23}} \\ H_1: \exists i, j \text{ tal que } \mu_i \neq \mu_j \end{cases}$$

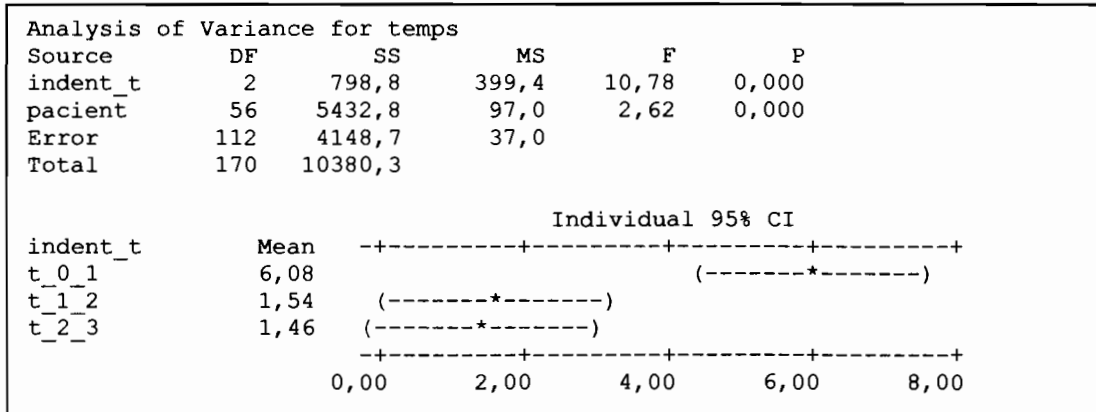


Fig.24. Test de comparació de les mitjanes dels temps entre arítmies.

Veiem que hi ha diferències significatives entre els temps entre arítmies (t_{01} , t_{12} i t_{23}), ja que tenim un p-valor de 0'000, amb el que rebutgem H_0 , d'igualtat entre mitjanes.

Estudi de l'additivitat del temps entre arítmies

A continuació presentem l'estudi de les sumes i les diferències del temps entre arítmies:

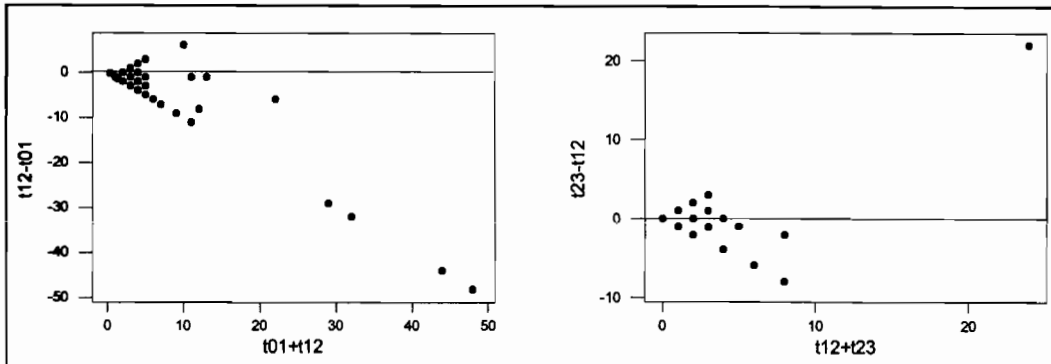


Fig.25. Plot de les sumes i les diferències dels temps entre arítmies.

Com en el primer estudi, en el gràfic de l'esquerra, veiem que la majoria de punts queden per sota de la línia d'equivalència, per tant temps es redueixen per a gairebé totes les observacions ja que $t_{12} < t_{01}$. En canvi, en el gràfic de la dreta, no es veu de forma tan aquesta situació.

Presentem ara la descriptiva:

Variable	N	N*	Mean	Median	TrMean	StDev
t12-t01	57	2	-4,54	-1,00	-2,86	9,98
t23-t12	57	2	-0,088	0,000	-0,196	3,527

Variable	SE Mean	Minimum	Maximum	Q1	Q3
t12-t01	1,32	-48,00	6,00	-4,00	-1,00
t23-t12	0,467	-8,000	22,000	0,000	0,000

Fig.26. Descriptiva de les diferències dels temps entre arítmies.

Veiem que la diferència entre t_{12} i t_{01} (-4'540) és més gran, en valor absolut, que la diferència que hi ha entre t_{23} i t_{12} (-0'088).

Com al primer estudi podem observar que en ambdues variables ($t_{12}-t_{01}$ i $t_{23}-t_{12}$) la variança és molt més gran que la seva mitjana (en valor absolut), prenent valors de $|-4'540| \ll 9'98^2$ i $|-0'088| \ll 3'527^2$ respectivament.

Cal comentar que, com hem vist abans, la dispersió augmenta amb el valor mig i per aquesta raó hem degut de fer l'anàlisi de les variables prenent logaritmes.

Realitzarem un test per tal de veure si la mitjana de les diferències entre els temps entre arítmies és zero o no:

$$\begin{cases} H_0 : \mu_{t_{21}-t_{01}} = 0 \\ H_1 : \mu_{t_{21}-t_{01}} \neq 0 \end{cases} \quad \begin{cases} H_0 : \mu_{t_{23}-t_{12}} = 0 \\ H_1 : \mu_{t_{23}-t_{12}} \neq 0 \end{cases}$$

Test of mu = 0 vs mu not = 0				
Variable	N	Mean	StDev	SE Mean
t12-t01	57	-4,54	9,98	1,32
t23-t12	57	-0,088	3,527	0,467
Variable	95,0% CI		T	P
t12-t01	(-7,19;	-1,89)	-3,43	0,001
t23-t12	(-1,024;	0,848)	-0,19	0,852

Fig.27. Prova d'hipòtesi d'una mitjana: $\mu=\mu_0$
per les diferències dels temps entre arítmies.

Observem resultats anàlegs: la diferència entre t_{12} i t_{01} és negativa i significativament diferent de zero, amb el que podem comprovar que hi ha acceleració. En canvi, la diferència entre t_{23} i t_{12} no és significativament diferent de zero. Així doncs, no observem acceleració.

Com es pot observar podem rebutjar que la mitjana de les diferències entre t_{01} i t_{12} ($\mu_{t_{21}-t_{01}}$) sigui zero, ja que tenim un p-valor de 0'001. En canvi, no hi han evidències per a rebutjar que la mitjana de les diferències entre t_{12} i t_{23} ($\mu_{t_{23}-t_{12}}$) sigui zero, ja que tenim un p-valor de 0'852.

També podem veure que, la diferència entre t_{12} i t_{01} té una mitjana de -4'54 (-7'190 ; -1'890), negativa i significativament diferent de zero. Per tant, el segon episodi triga en aparèixer entre 2 i 7 mesos menys que el primer, amb el que podem comprovar que el temps es redueix.

En canvi, la diferència entre t_{23} - t_{12} té una mitjana de -0'088 (-1,024; 0,848). Veiem que el tercer episodi triga en aparèixer entre 1 mes menys i 0'8 mes que el segon, l'interval inclou el zero. Així doncs, no observem cap reducció de temps.

Cal comentar que els resultats obtinguts són els mateixos que a l'estudi anterior.

A continuació presentem el test de normalitat de per tal de comprovar que les $t_{12} - t_{01}$ compleix amb les premisses de normalitat:

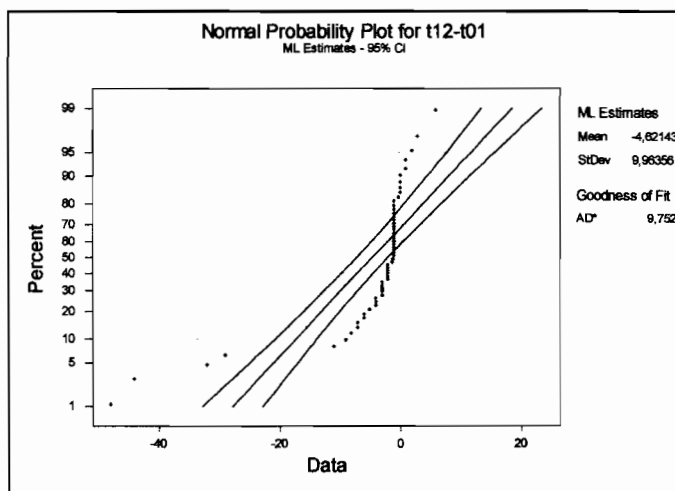


Fig.28. Gràfic de normalitat del model additiu

Com es pot veure al gràfic de normalitat del estudi de l'additivitat del temps entre arítmies, la diferència del temps t_{12} i t_{01} no segueix una normal.

Estudi de l'additivitat en escala logarítmica del temps entre arítmies

Com hem comentat a l'anàlisi anterior, la varianza de la diferència dels temps entre arítmies és molt més gran que la mitjana i la dispersió augmenta amb el valor mig. Per un altre lloc, hem vist que el gràfic de normalitat de les sumes i les diferències dels temps entre arítmies no s'ajusta bé a una normal. A continuació realitzem transformacions logarítmiques per tal de resoldre aquestes qüestions.

Presentem l'estudi fent les sumes i les diferències dels logaritmes dels temps entre arítmies:

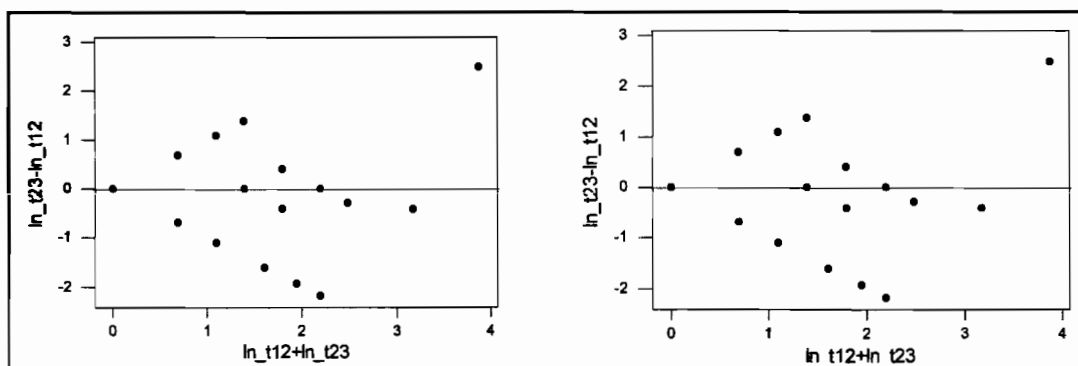


Fig.29. Plot de les sumes i les diferències dels temps entre arítmies.

Es repeteix la situació anterior, la majoria de punts queden per sota de la recta d'equivalència ($t_{12} < t_{01}$) al gràfic de l'esquerra. Al gràfic de la dreta no es veu tan clara aquesta situació. Igualment en ambdós casos podem dir que, a mida que el temps es va fent més gran la divergència augmenta: heterocedasticitat.

Presentem la descriptiva dels logaritmes de les diferències:

Variable	N	N*	Mean	Median	TrMean	StDev
ln_t12- ln_t01	57	2	-0,968	-0,693	-0,916	1,056
ln_t23- ln_t12	57	2	-0,0908	0,0000	-0,0746	0,7273

Variable	SE Mean	Minimum	Maximum	Q1	Q3
ln_t12- ln_t01	0,140	-3,892	1,099	-1,386	-0,347
ln_t23- ln_t12	0,0963	-2,1972	2,4849	0,0000	0,0000

Fig.30. Descriptiva de les diferències dels logaritmes dels temps entre aritmies.

Veiem que la diferència entre $\ln t_{12}$ i $\ln t_{01}$ (-0'968) és més gran, en valor absolut, que la diferència entre $\ln t_{23}$ i $\ln t_{12}$ (-0'0908).

Pel que fa a la varianza de $\ln t_{12} - \ln t_{01}$ resulta més petita que la mitjana (en valor absolut), prenent les següents valors $1'056^2 < |-0,968|$ respectivament i pel que fa a la varianza de $\ln t_{23} - \ln t_{12}$ continua sent més gran que la mitjana $0'7273^2 > |-0'0908|$, com a l'estudi anterior.

Realitzem un test per veure si les diferències dels logaritmes dels temps són significativament diferents de zero o no:

$$\left\{ \begin{array}{l} H_0 : \mu_{\ln t_{21} - \ln t_{01}} = 0 \\ H_1 : \mu_{\ln t_{21} - \ln t_{01}} \neq 0 \end{array} \right\} \left\{ \begin{array}{l} H_0 : \mu_{\ln t_{23} - \ln t_{12}} = 0 \\ H_1 : \mu_{\ln t_{23} - \ln t_{12}} \neq 0 \end{array} \right.$$

Test of mu = 0 vs mu not = 0				
Variable	N	Mean	StDev	SE Mean
ln_t12-ln_t01	57	-0,968	1,056	0,140
ln_t23-ln_t12	57	-0,0908	0,7273	0,0963

Variable	95,0% CI	T	P
ln_t12-ln_t01	(-1,249; -0,688)	-6,93	0,000
ln_t23-ln_t12	(-0,2838; 0,1022)	-0,94	0,350

Fig.31. Prova d'hipòtesi d'una mitjana: $\mu = \mu_0$ per les diferències dels logaritmes dels temps entre aritmies.

Observem que la primera mitjana de la diferència de logaritmes ($\mu_{\ln t_{21}-\ln t_{01}}$) és $-0'968$, negativa i significativament diferent de zero, ja que el p-valor és zero (podem rebutjar la hipòtesi nul·la). En els intervals de confiança, veiem que el segon episodi triga en aparèixer entre 28'679% ($e^{-1'249}$) i un 50'026% ($e^{-0'688}$) menys que el primer. Amb el que podem comprovar una reducció del temps.

En canvi, la segona mitjana ($\mu_{\ln t_{23}-\ln t_{12}}$) és $-0'0908$ ($-0'2838$; $0'1022$) no significativament diferent de zero, ja que el p-valor és 0'0350 (no tenim evidències per rebutjar H_0). A més a més el seu interval de confiança inclou el zero. Per tant podem dir que, no hi ha reducció del temps a partir de la segona arítmia.

A continuació presentem el test de normalitat de per tal de comprovar que les $t_{12} - t_{01}$ compleix amb les premisses de normalitat:

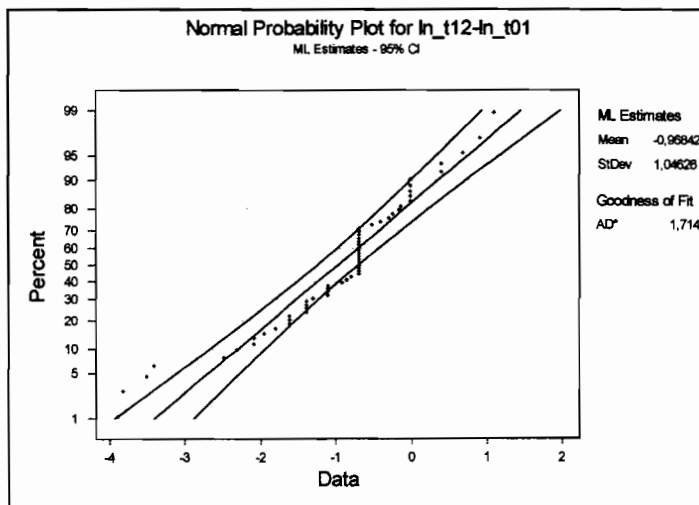


Fig.32. Gràfic de normalitat del model multiplicatiu

Com es pot veure, tots els punt cauen dins dels intervals, per tant podem dir **que el gràfic de normalitat surt molt millor en aquest estudi de l'additivitat en escala logarítmica** que amb l'estudi anterior d'additivitat del temps entre arítmies. Pel que **triariem aquest model** per explicar l'evolució del temps entre arítmies.

Com hem pogut veure al llarg de tot aquest estudi, hem arribat a les mateixes conclusions: el temps lliure d'arítmies disminueix.

ANNEX 5: Programa SAS Simulacont

```

%macro NP;
53
%mend NP;

%macro TT;
56
%mend TT;

%macro L;
1 5
%mend L;

data simulacont;

do i=1 to %NP;
    compt = 0;
    x = ranexp (%L);
    do while (x < %TT);
        compt = compt + 1;
        y = ranexp (%L);
        x = x + y;
    end;
    numepic = compt;
    output;
end;
run;

```

ANNEX 6: Programa SAS Simuladisc

```

%macro PI;
1 5
%mend PI;

%macro ITER;
112
%mend ITER;

%macro NP;
53
%mend NP;

data simuladisc;

do i=1 to %NP;
    compt = 0;
    do j=1 to %ITER;
        u = uniform(0);
        if (u < %PI) then compt = compt + 1;
    end;
    numepid = compt;
    output;
end;
run;

```

ANNEX 7: Col·laboració en la tesi doctoral de Joaquina Belchi

Abans de començar el nostre projecte, vam fer l'anàlisi estadístic per a la tesi doctoral de Joaquina Belchi.

La realización del estudio estadístico de la tesis doctoral con el título: "Las arritmias auriculares en los pacientes portadores de desfibrilador automático implantable. Ventajas de los sistemas bicamerales para su estudio, control y tratamiento", que fue depositada y aprobada y en la actualidad tras haber aprobado el tribunal (número de registro 2661 del Centro: Facultad de Medicina y Odontología) está pendiente de lectura, fue realizada por Raquel López y Fabiana Montes de Oca.

Quiero mostrar especial atención a la capacidad de ambas en interpretar lo que el investigador les solicitaba así como el interés y rapidez con que han realizado este trabajo, sin cuya ayuda no hubiera sido posible su realización.

Fdo: Joaquina Belchí Navarro

La nostra feina va consistir en respondre, seguint diferents mètodes estadístics, a les preguntes que li van sorgir per tal d'extreure les màximes conclusions partint de les dades que va recollir dels seus pacients. Cal recordar, que existeixen dues poblacions: els que tenen la indicació del DAI per TV/FV i els que la tenen per FA.

Els seus objectius es poden resumir en:

- Veure les característiques de totes les arítmies en les dues poblacions. En concret, per auriculars s'ha de veure l'eficàcia del tractament.
- Jutjar les capacitats de diagnòstic del dispositiu a l'hora d'interpretar un o altre tipus d'arítmia.

Cal fer alguns aclariments:

- Dins de les arítmies auriculars hi ha 2 tipus: TA i FA (s'analitzaran juntes i per separat). Juntes quan es vulgui veure la relació entre les arítmies auriculars i les ventriculars i per separat quan vulguem veure la descriptiva i l'eficàcia del diagnòstic i del tractament. Pel que fa a les arítmies ventriculars, existeixen varis tipus: TV, FV, doble taquicàrdia
- Quan a la variable electrograma apareix ND, significa que l'arítmia he estat tan curta que no ha donat temps a que es registri a l'electro. Sempre fa referència a arítmies auriculars i s'anomenen 'no sostingudes'. A la resta d'arítmies auriculars les anomenarem sostingudes. Per tant, dins de les arítmies auriculars tindrem, segons el tipus, TA o FA i sostingudes i no sostingudes, segons si disposem o no de l'electrograma.

Alguns dels estudis que ens han demanat són:

1. Descriptiva: tot s'ha d'analitzar de forma global (les dues poblacions juntes) i per separat, segons els tipus de població al que s'implanta el dispositiu.
 - 1.1. Quantes arítmies de cada tipus s'han presentat? Fer-ho segons la variable electrograma ja que és el que l'investigador analitza un cop ha vist l'electro. Nombre mig de recurrències d'arítmies auriculars en general.
 - 1.2. Per a cada tipus d'arítmia (TA, FA, TV, FV, doble taquicàrdia):
 - Hora d'inici
 - Duració
 - Longitud de cicle
 - 1.3. Existeixen diferències entre la longitud del cicle les arítmies sostingudes i de les no sostingudes? També cal saber si la longitud

- del cicle es relaciona amb els fàrmacs que s'estan prenent i si el fet de tenir molts episodis d'arítmies auriculars no sostingudes està relacionat amb el fet de què tingui més episodis de sostinguts.
- 1.4. Pel que fa a la hora d'inici, s'ha de separar entre TA sostinguda i TA no sostinguda i entre FA sostinguda i FA no sostinguda, per veure si les no sostingudes s'inicien a hores diferents de les sostingudes.
 - 1.5. Quants pacients han presentat TA, quants FA i quants ambdues arítmies auriculars (TA+FA). Per una altra banda, quants han presentat TV, FV o les dues. Es correlaciona el fet de tenir més episodis auriculars amb tenir més de ventriculars?
 - 1.6. Existeixen variable clíniques que prediguin el tenir més arítmies de TA/FA? Es relaciona TA o FA amb alguna cardiopatia base?
 - 1.7. Hi ha una variable recodificada que classifica si ha estat el primer episodi o no (*idprep*). Interessa saber el temps mig des de l'implant fins a l'aparició del primer episodi. Diferenciar entre les dues poblacions i veure si existeixen diferències.
 - 1.8. Els pacients que tenen activats el *swbd* i *are*, tenen un menor nombre d'arítmies auriculars? Es relaciona el mode d'estimulació (*modest*) amb tenir més o menys arítmies auriculars?
 - 1.9. Quan existeix diferència entre el que es detecta a l'inici (*detecció inicial*) i a la teràpia, es prediu un major nombre de transicions en l'evolució de l'episodi? Es prediu una major ineficàcia del tractament?
 - 1.10. Hi ha diferències, segons si la programació està en ON o en OFF, per a la proporció d'arítmies (TA o FA), la longitud del cicle, la distribució horària, burden (temps en FA o en TA respecte el temps total de seguiment)?
 - 1.11. Disminueix la freqüència d'episodis auriculars al llarg del seguiment?
 - 1.12. Pels episodis de doble taquicàrdia:
 - % del total de les arítmies ventriculars.
 - % de pacients amb almenys 1 episodi de taquicàrdia dual.
 - % de TV i % de FV (*detecció inicial*)
 - Quin tipus d'arítmia auricular és documentada a l'inici (*rittv*)?
 - La durada de l'arítmia auricular prediu que hi hagi doble taquicàrdia?

- % de malalts que després passen a un ritme sinusal i quants queden en FA o TA. Dels casos que surtin en FA o TA, quin temps fins al següent episodi d'arítmia ventricular és menor?

1.13. Del grup de població d'arítmies auriculars, quin percentatge té ventriculars? Es relaciona la seva presència amb alguna variable clínica?

2. El següent anàlisi es veure la capacitat de diagnòstic del dispositiu.

Es tracta de veure quantes vegades el dispositiu ha detectat arítmies ventriculars (TV, FV, Doble Taquicàrdia) i l'investigador a l'analitzar l'electro ve que són auriculars o una altra cosa, com per exemple, l'artefacte, sobre detecció de T... Per fer-ho s'han de comparar les variables *detecció preteràpia* i *electrograma*. La següents variables fan referència a paràmetres que programa l'investigador al dispositiu per a millorar la detecció (*sv*, *prlogic*, *ts*, *tafa*, *tpsv*). Existeix relació entre que estiguin activats i l'exactitud del diagnòstic? En els casos que es detecti de forma inadequada, es relaciona la longitud del cicle amb el *svt-i*?

Per una altra banda, s'ha de veure la capacitat del dispositiu de diferenciar dins de les arítmies auriculars si és Ta o FA. Per fer-ho, s'han de compara les variables *detecció preteràpia* i *electrograma* (aquí els episodis no sostinguts no compten perquè l'investigador no té cap element per fer el judici ja que no disposa de l'electro).

3. S'ha de veure l'eficàcia del tractament (només del episodis amb detecció adequada).

3.1. Descriptiva general del nombre total de Rampas, Burst, Hz i Xocs que s'han administrat per TA+FA, TA i FA. També separar-ho segons la població. Amb la resta d'arítmies i de teràpies igual.

3.2. Posteriorment s'ha de veure el resultat de la teràpia (pot ser cese, *no cese*, FA, TA, FV, TV. Això és degut a que a l'administrar un tractament l'arítmia pot transformar-se en una altra). Amb tot això, s'ha de quantificar el resultat de cada tipus de teràpia, és a dir, quants episodis de cada tipus han cessat en cada teràpia. Per exemple, 60 episodis de TA van cedir amb 6 seqüències de Burst.

Això amb totes les teràpies i amb totes les arítmies i, a més, diferenciar-ho posteriorment segons el tipus de població.

L'anàlisi que farem a continuació amb aquestes dades serà el següent:

- 3.2.1. Quants episodis de TA, FA, TV, FV, ... van finalitzar amb èxit en cada tipus de teràpia? Quin tractament és més eficaç per a cada tipus d'arítmia? Veure l'eficàcia global de les teràpies d'estimulació (Burst, Rampes, Hz juntes) per el tractament d'arítmies auriculars i després per separat. És més eficaç en alguna de les dues poblacions?
- 3.2.2. Quants episodis van necessitar més d'un tipus de teràpia pel seu cas?
- 3.2.3. Quants episodis s'han transformat en altre tipus d'arítmia després de l'administració d'una teràpia? Quants episodis de TA o FA després de la teràpia s'han transformat en TV o FV? Alguna transformació ha estat més freqüent en algun tipus de teràpia?
- 3.2.4. Quan una arítmia al electrograma és FA i després com a conseqüència de les teràpies es transforma en TA, és més eficaç el Xoc quan s'administra que en el cas de què continuï en FA?
- 3.2.5. Es relaciona la eficàcia del tractament amb la longitud del cicle? I amb els fàrmacs que s'estan prenent?
- 3.2.6. Quants episodis van finalitzar en el primer minut?
- 3.2.7. Quin va ser el nombre mig d'episodis per pacient?
- 3.2.8. Influeix el retard en l'estimulació en l'èxit o fracàs de la teràpia?
- 3.2.9. Compara la durada del episodi quan cessa amb estimulació amb la durada del episodi en el què falla l'estimulació.
- 3.2.10. Influeix el retard en els Xocs en la seva eficàcia?
- 3.2.11. Es relaciona algun tipus de teràpia o algun tipus de cardiopatia amb alguna forma de finalització (*resulfin*)? Fer una descriptiva de la forma de finalització dels episodis. En els episodis que finalitzen de forma espontània, quin és el patró de finalització?
- 3.2.12. De què depèn més l'èxit, de que sigui TA o FA o de la longitud de cicle?

4. Anàlisi de la capacitat que té el dispositiu.

Veure quantes vegades el dispositiu detecta finalització (*findat*) i l'investigador detecta que no ha cessat (*interdat*).